Under the Word Shower:

Massive Repetition Priming of Words and Pseudowords

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Thorsten Albrecht
aus Braunschweig

1. Referent:	Professor Dr. Dirk Vorberg
2. Referent:	Professor Dr. Reinhold Kliegl
eingereicht am:	08.10.2007
mündliche Prüfung (Disputation) am:	10.03.2008

Druckjahr 2008

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Tagungsbeiträge

Albrecht, T. & Vorberg, D. (2006). Frequenzeffekte unter der Wortdusche. In H. Hecht, S. Berti, G. Meinhardt & M. Camer (Hrsg.), *Experimentelle Psychologie. Beiträge zur 48. Tagung experimentell arbeitender Psychologen, Johannes-Gutenberg-Universität Mainz,* S.157.

Albrecht, T. & Vorberg, D. (2007). „Unter der Wortdusche“: Ein neues RSVP-Paradigma zur Untersuchung von Priming-Effekten. In K.F. Wender, S. Mecklenbräuker, G. D. Rey & T. Wehr (Hrsg.), *Experimentelle Psychologie. Beiträge zur 49. Tagung experimentell arbeitender Psychologen, Universität Trier,* S.196.

SUMMARY **5**

1. INTRODUCTION **6**

1.1 LONG-TERM REPETITION PRIMING 8
1.2 LONG-TERM REPETITION PRIMING OF NONWORDS 12
1.3 SHORT-TERM REPETITION PRIMING 15
1.4 ABSTRACTIONIST VS. EPISODIC THEORIES ON PRIMING AND WORD RECOGNITION 18
1.4.1 Multiple Read-Out Model (MROM, Grainger & Jacobs, 1996) *19*
1.4.2 Retrieving effectively from memory – lexical decision (REM-LD) *20*
1.5 NEURONAL CORRELATES OF REPETITION PRIMING 22
1.6 OPEN QUESTIONS 25
1.7 THE ACCOUNT OF THIS STUDY 26
1.7.1 Data analysis and the "language-as-fixed-effects-fallacy" *30*
1.8 GENERAL METHOD 35

2. EXPERIMENT 1A AND 1B: BASICS AND DETAILS OF THE METHOD **39**

2.1 EXPERIMENT 1A 39
2.1.1 Methods and Materials *39*
2.1.2 Results *42*
2.1.3 Discussion *47*
2.2 EXPERIMENT 1B 50
2.2.1 Methods *50*
2.2.2 Results *50*
2.2.3 Discussion *55*

3. EXPERIMENT 2: MANIPULATING CONSCIOUS RECOLLECTION **57**

3.1 METHODS 57
3.2 RESULTS 58
3.3 DISCUSSION 60

4. EXPERIMENT 3: QUANTIFYING RECOGNITION OF REPEATED PRIMES **61**

4.1 METHODS 62
4.2 RESULTS 63
4.3 DISCUSSION 67

5. EXPERIMENT 4: THE LONGEVITY OF MASSIVE PRIMING **69**

5.1 METHODS 69
5.2 RESULTS 70
5.3 DISCUSSION 71

6. EXPERIMENT 5: CROSS-PRIMING OF WORDS AND PSEUDOWORDS **73**

6.1 METHODS 73
6.2 RESULTS 74
6.3 DISCUSSION 75

7. EXPERIMENT 6A AND 6B: SHORT STREAMS AND LONG LAGS **77**

7.1 EXPERIMENT 6A 77
7.1.1 Methods and Materials *78*
7.1.2 Results *83*
7.1.3 Discussion *87*
7.2 EXPERIMENT 6B 89
7.2.1 Methods and Materials *89*
7.2.2 Results *92*
7.2.3 Discussion *94*

8. GENERAL DISCUSSION **96**

9. REFERENCES **104**

Summary

How does a "word shower" of more than 200 words rushing by with a speed of approximately 80 ms per item influence the later processing of the words, e.g. in a lexical decision task? Does presentation of single words and pseudowords have any impact on later processing despite the long time span between presentation and test, and the high amount of distractors present in the word stream? If so, does a higher presentation frequency cause effects to cumulate?

I presented streams of words and pseudowords during which participants had to count forenames occurring in the stream. Words and pseudowords occurred with various presentation frequencies. Immediately after the word stream participants performed a series of lexical decisions on words and pseudowords, which had either been presented in the preceding stream or not. In Experiments 1a and 1b massive repetition priming in word streams was assessed with single presentation durations of 150 ms and 56 ms, respectively. The relation of conscious perception and later recollection of single items and priming was assessed in Experiments 2 and 3. The dependence of priming effects of the delay between presentation and test showed no decrease of priming over several minutes (Experiments 1a, 1b, and 4). Further, cross-priming of pseudoword neighbours on words and vice versa (Experiment 5) showed reduced pseudoword-to-word priming but normal word-to-pseudoword priming The last chapter focussed on a refinement of the method by shortening the word streams and including only one critical item per word stream. Long-lag priming over one day was compared for response relevant primes and non-relevant primes (Experiment 6a) highlighting the importance of response learning. In Experiment 6b recognition and priming was assessed in the same trial, and advantages and limitations of this short word stream paradigm are discussed.

In sum, the experiments demonstrated small but reliable and long-lasting cumulative repetition priming effects that were positive for words, but negative for pseudowords indicating a growing familiarity or "word-likeness" of pseudowords. Crucially, recognition performance seemed not to be causal for the effect to occur. Thus, the "word shower" proved to be a paradigm suitable to induce long-term priming effects by short but massive prime presentation.

1. Introduction

In every moment a continuous stream of information reaches our senses and the nervous system, but only a small fraction of this stream comes to awareness, and a great deal of information is lost once it has gained awareness. Nevertheless it is widely acknowledged that stimuli that did not reach consciousness or that cannot be recollected at the time of test can influence cognitive processes and behaviour. This has been shown extensively in studies on implicit learning, implicit memory, and in various priming paradigms. However, most studies presented stimuli in relative isolation, i.e., the critical stimulus and the test stimulus were the only stimuli to be processed by the central nervous system. Such studies lack ecological validity in that they provide evidence for the possibility of influence, but do not give any information about the extent of such influence in a “real-world” situation. The question is whether the presentation of one stimulus has an influence on the processing of a later episode when it is presented in a “continuous” stream of non-relevant stimuli, in which processing of single items is degraded. For example, imagine a “word shower” of several hundreds of words rushing by with a speed of 60 ms per item. What impact does this “shower” have on a later task?

If one assumes that words are represented at an abstract lexical level in the human brain (i.e., *the mental lexicon*), and that, on presentation of a specific word, the lexical representation of this word becomes activated in some sense, then it is reasonable to further assume that residual activation of lexical word representations may influence behaviour in a later situation. The question is, whether activation is an all-or-nothing process (i.e., inactive state vs. active state) or a cumulative process (i.e., continuous increase from inactivity to full activity). An all-or-nothing process would resemble some kind of memory buffer, which either contains a word representation (active) or it does not (inactive). In the active state a word representation may influence cognitive processes to a certain degree, in the inactive state it does not. A cumulative process of activation allows for different levels of activation and, therefore, a word representation in a less activated state should influence cognitive processes to a smaller degree, than a word representation in a more active state. In addition, in an all-or-nothing activation process the activation is the same regardless of whether a word is presented once or twice in rapid succession. A cumulative account predicts a greater amount of activation when a word is presented twice than when it is presented once. So, imagine a “word shower”, in which words are presented various times: According to an all-or-nothing account the presentation of a word shower should influence later behaviour regardless of presentation frequency of individual words, whereas an cumulative account predicts greater impact of words presented more often in a “word shower”. Thus, the “word shower” paradigm can be useful in investigating the representation of words in the human brain and the structure of a mental lexicon by presenting various words in rapid succession with different presentation frequencies and assessing the impact of these words on the behaviour in a later situation.

One paradigm which is particularly suitable to exemplify the lasting impact of one episode on a later situation is repetition priming of words (e.g. Scarborough, Cortese & Scarborough, 1977): The encounter of a word (the prime) influences the behaviour on a second encounter of the same word. In a typical repetition priming experiment words are presented in a study task and again later in a test task. For example, participants are asked to read the following list of words aloud first: GRAPE, CURVE, DOG, HOUSE, BATH. On a second task, these same words, new words (e.g. PAPER, CANDLE, BANK, FLOWER, OVEN), and nonwords (e.g. GREACH, STROFE, SOM, BONT, CANG) are presented singly one by one. On each presented item participants decided whether the presented item is a word or a nonword. Repetition priming describes the phenomenon that the response to a word that is presented the second time (e.g. CURVE) is facilitated in that participants can react faster and more accurate to it than to words that had not occurred before (e.g. CANDLE). In the present study I investigated whether repetition priming occurs, when items are presented repeatedly in a “word shower” of several hundred items, and are tested in a later lexical decision task. To what extent is priming cumulative beyond the first repetition, i.e., are the facilitatory effects for words exaggerated when the word is presented 5 times or 10 times within one “word shower”?

Repetition priming – also termed identity priming (Hutchison, Neely, Neill & Walker, 2004) or direct priming (Graf & Schacter, 1985) – has been studied extensively in several fields of psychological research, e.g. word recognition, implicit memory research, (unconscious) perception and word acquisition. In general one distinguishes between two broad classes of paradigms used in repetition priming studies, which have led to different results and accounts: Long-term paradigms yield effects that last from minutes to days, which have been explained mostly by *episodic accounts*. In general, it is assumed that on presentation of a word some form of episodic memory is constructed or that an existing memory is enriched. On the next presentation this enriched memory trace is retrieved more easily than on the first encounter. Short-term paradigms yield effects that decay within seconds to minutes, and have given rise to abstractionists or lexical accounts of repetition priming. It is assumed that the process of word recognition involves the activation of an abstract representation of the presented word. This activity resides for some time in the neural network, and leads priming effects on a second presentation. In the following sections I will describe the different paradigms in detail, discuss the episodic and abstractionist approach, and describe two models that can account for word recognition and repetition priming. Eventually I will review findings on the neural correlates of repetition priming. In each section I will focus on the question outlined above: To what extent are repetition effects relevant in a "real world-like" scenario, i.e., in a stream of (almost) continuous information flow.

1.1 Long-term repetition priming

Long-term repetition priming refers to priming effects that last from minutes up to several days or even years without or with little decay only. Most studies that obtained long-term effects have employed an experimental setup in which a word occurred repeatedly in the course of the experiment, and participants performed a response on each occurrence of this word. This may be the same task on each occurrence, or one study task and a different test task. For example, participants may perform continuously lexical decisions on single words and nonwords with repeating occurrences of items at varying lag (same task), or participants may get a list of words to memorize at study, and some time later perform a lexical decision on learned and new words at test (different tasks). Long-term priming has been shown with various stimulus types, e.g. objects (van Tourenout, Ellmore & Martin, 2000; Cave, 1997), words (Jacoby, 1983), and nonwords (Wagenmakers, Zeelenberg, Steyvers, Shiffrin & Raaijmakers, 2004; Stark & McClelland, 2000, Bentin, 1989). Typical tasks used in long-term paradigms are e.g. naming (Cave, 1997), speeded word reading (MacLeod & Masson, 2000), word fragment completion (Graf & Schacter, 1985), identification (Jacoby, 1983) or lexical decision (McKone, 1998, 1995; Ratcliff, Hookley & McKoon, 1985; Scarborough et al., 1977). It has been argued that long-term effects crucially depend on participants' awareness of the first presentation (Forster & Davis, 1984), or first and second presentations have to be perceived as separate entities in time (Tenpenny, 1995).

Long-term paradigms are employed mostly in implicit memory research. The focus here is not on unconscious and automatic encoding processes, since the target is fully visible and processed on each occurrence, but on automatic retrieval. Implicit memory is defined as memory that enhances performance but does not depend on conscious recollection (Schacter, 1987). Several researchers have demonstrated a dissociation of explicit and implicit memory in that implicit memory influenced the behaviour of participants and patients in the absence of explicit memory (Graf & Schacter, 1985; Tulving, Schacter & Stark, 1982). But only a few studies on repetition priming in healthy participants included tests of conscious episodic memory, because the tasks used are assumed to measure true implicit effects. Ratcliff et al. (1985) contrasted pure repetition effects with recognition effects: They continuously presented words and nonwords, on which participants either performed a recognition task or a lexical decision. Repetitions of the same item occurred at varying lags of up to 16 intervening items. Recognition performance, measured by response times and accuracy, steadily decreased with increasing lag. Facilitation on lexical decisions sharply decreased with increasing lag, but remained at a constant level for lags beyond two intervening items. The authors concluded that repetition priming in lexical decision is based on two mechanisms, an intermediate component that is also part of recognition and a long lasting repetition component. In a similar study McKone (1995) claimed to have identified an additional "short-term" component of repetition priming that overlaid the traditional long-term component in lexical decision. This short-term priming component is said to decay in up to several seconds, and follows a different time course for words and for nonwords. McKone argues that this component is mediated by existing traces in the mental lexicon, and that the reduction of this effect is due to decay over time and to interference from intervening items (McKone, 1998). However, she did not control for recognition processes, so the fast decaying short-term component may be largely due to explicit recognition, whereas the smaller long-term component is due to repetition, as Ratcliff et al. (1985) pointed out.

Another main finding in repetition priming of words is the lexicality effect, the stable phenomenon that repetition priming is stronger for low frequency words (i.e., words that rarely occur in a language) than for high frequency words (e.g. Wagenmakers et al., 2004, Stark & McClelland, 2000, Scarborough et al, 1977). This effect usually is interpreted as evidence for a lexical/semantic locus of repetition priming of words. However, it has been widely shown that the amount of repetition priming is influenced by the contextual similarity of study and test phase (Jacoby, 1983; Masson & MacLeod, 1992, Masson & MacLeod, 2002, Stark & McClelland, 2000; Bowers, 2000b), which points to an episodic locus of repetition priming. For detailed reviews of episodic vs. lexical accounts, see Tenpenny (1995) and Bowers (2000a).

In most studies on long-term repetition priming the items were presented twice, once in the study task and once again at test. Feustel et al. (1983) employed a continuous identification task, in which targets occurred in up to four trials. Targets either were words or “pseudowords”, i.e., letter strings that suffice the grammar and pronunciation rules of a given language as e.g. “BOCH”. Identification performance (measured as response time) increased for words and pseudowords with increasing repetition number, i.e., beyond the second presentation. In a more recent study Wagenmakers et al. (2004) obtained similar results in a continuous lexical decision task. Rueckl (1990) instructed participants to read aloud lists of words once or three times, and in a test phase participants had to identify the same words, orthographically similar words, and new words in a masked identification task. He obtained a stable priming effect for both repeated words and similar words, but a further benefit from words presented three times was obtained only for the same words, but not for similar words.

Masson and Hicks (1999) investigated whether non-targets embedded in a sequence of rapidly presented words would exhibit long-term repetition priming effects, and if the selection of an item for an overt response at study influences the magnitude of priming. In the study phase they presented sequences of six rapidly presented words with one target word and one critical word at two successive positions within the list. Lists were presented at rates from 200ms per word up to 1000 ms/word. Before each list presentation, participants were given a category name (e.g. COLOUR). During list presentation participants had to search for a word from this category. Half of the lists contained such a target word together with a critical word preceding the target word. For example, given the category COLOUR, and the list *sung, bath, pink, curb, dawn, junk*, "bath" would be the critical word, and "pink" would be the target word. In the test phase the critical words were presented in a word identification task. Regardless of presentation rate of the lists a small repetition priming effect was observed, although the critical word never had been search targets before. This repetition effect was enlarged, when participants either read the list items aloud at study or the search targets at study were to be identified at test. However, if the same search task was employed at study and at test, no repetition priming was observed unless the search targets were the same at study and at test. Masson and Hicks (1999) concluded that response-based encoding is a fundamental principle of repetition priming mechanisms, and that full perception of the target word is necessary for repetition priming to occur. They strongly argue against abstractionist accounts, because repetition priming on the basis of lexical activation does not predict any context or encoding effects. However, the encoding processes which led to the disappearance of priming effects in the Masson and Hicks study may work on a different (response-based) level than repetition priming itself: The word lists used at study were short (6 items) and had a "label" each (the semantic category for the search task), by which participants monitored them. Thus, each list constituted a distinct entity in time and content. In masked word identification this may not play a role, because orthographical information is sufficient for performing the task. In the search task at test, however, the same lists were presented as in the study phase either with only the search category changed (same context) or with the search category and the critical item changed (different context). So, a test list at least had 5 out of 6 items in common with the corresponding study list. Under the assumption that lists are encoded as distinct entities with a specified response category at study, this should lead to a conflict at test diminishing any priming effects. Further, presentation durations of more than 200 ms together with the brevity of lists makes it possible for participants to categorize each item of a list before responding to the search task, so that participants might have responded to each single item even if not reporting so.

1.2 Long-term repetition priming of nonwords

Findings on long-term repetition priming for pseudowords and nonwords are taken as arguments in for an episodic approach because no lexical entries for nonwords exist in the mental lexicon and therefore repetition priming solely based on lexical mechanisms cannot occur at all. But overall, evidence for repetition priming of pseudowords and nonwords is mixed.

In a series of experiments Stark and McClelland (2000) investigated repetition priming for words, pseudowords and nonwords in a continuous identification task with recognition (CID-R). In a study task participants had to judge how much they did like singly presented 4-letter strings. In the test task, on each trial a letter string was presented progressively demasked until participants identified it. After identification they reported whether the actual letter string had already occurred in the study phase (‘old item’) or not (‘new item’). With this procedure Stark and McClelland found reliable repetition effects for words and pseudowords, which in general were greater for pseudowords. Nonwords showed small but reliable repetition effects, when they comprised vowels and consonants, but not when letter strings consisted of consonants only. In addition, the results gave evidence for a slight contamination of repetition effects by explicit retrieval processes, but nevertheless, there still was stable priming for words and pseudowords, which could not be explicitly recollected. McKone and Trynes (1999) also reported positive repetition priming effects for pseudowords and nonwords in a same-different decision task, which did not differ from priming effects for words in amount.

For lexical decision tasks the evidence for nonword repetition priming is less clear. There are incosistent findings of facilitation, inhibition or no effects (for a review see Tenpenny, 1995). As pointed out by several researchers (Bowers, 1994; Feustel et al, 1983, Logan, 1988, Tenpenny, 1995) effects are facilitatory when lexical decision is used as study and test task, but not when used as test task only. In a more recent study Wagenmakers et al. (2004) varied the speed stress under which participants had to make lexical decisions on words and pseudowords. They employed a continuous lexical decision paradigm, in which participants performed several blocks of lexical decisions. Each target was presented in each of five blocks, so that each target was presented five times over the whole experiment. When participants received a “respond-when-ready” instruction, priming effects were facilitatory for words and for nonwords as measured by accuracy and response times. In contrast, when participants received a “signal-to-response” instruction the results differed substantially. For words, repetition priming was facilitatory again, but for pseudowords repetition inhibition was observed. However, this inhibitory effect occurred in the accuracy data only, but not in response times. On basis of these findings the authors argue in favour of two distinct processes that mediate repetition priming for pseudowords, one fast inhibitory process, based on familiarity of an item, and one facilitatory process, based on episodic memory traces of an earlier encounter. In another study Zeelenberg, Wagenmakers and Shiffrin (2004) corroborated the earlier findings that task characteristics at study (Experiment 1: lexical decision vs. letter height decision) and instructional focus at test (Experiment 2: accuracy vs. speed) can influence the direction of pseudoword repetition priming. Unfortunately, they did not assess explicit memory retrieval. Explicit memory for pseudowords may differ with encoding conditions.

Confirmation for a two process account for lexical decision of pseudowords comes from Perea, Rosa and Gómez (2005). They created pseudowords by replacing one letter from either low frequency words or high frequency words and had their participants perform lexical decisions on these items. For analysis the authors identified the response time, which was faster than 90 % of all responses (0.1-quantile) for each participant and condition. An ANOVA on these quantile data showed that responses on low frequency pseudowords were faster than on high frequency pseudowords. An analysis conducted with the mean response times (0.5-quantile) did not show any reliable difference. The authors interpret these results as evidence for a two process approach of lexical decision in pseudowords, one fast process of lexical activation, and one slow verification process.

None of the studies mentioned so far allow inferences whether the mechanism of long term priming is lexical or nonlexical in nature. In general, all findings can be interpreted in lexical/abstractionists view or in episodic retrieval view, although there is a hot debate about the mechanisms (see Tenpenny, 1995; Bowers, 2000a, for a review). However, some studies from outside the field of repetition priming gave evidence that nonwords develop lexical characteristics when repeatedly processed. One of the key concepts of abstractionist models of word recognition is lexical competition (Grainger & Jacobs, 1996, McClelland & Rumelhart, 1981), i.e., simultaneous activated word representations compete for recognition. Bowers, Davis and Hanley (2005) showed that pseudowords can take part in lexical competition. They constructed pseudowords by replacing one letter of words that did not have any so called "neighbours" (i.e., a word that differs in only one letter, e.g. "KETTE" and "KUTTE"). The pseudowords were repeatedly typed and read by participants in the study phase. In the test phase they performed a semantic categorization task (natural vs. artefact) on words that either were the parents of the studied pseudowords or not. Test phases were conducted immediately after study, one day later and one day later after another study phase. By employing a semantic categorization task on stimuli that were not studied before, the authors minimized possible influences of episodic memory traces. The results were clear cut: words whose pseudoword neighbours were studied, were consistently categorized slower than words without pseudoword neighbour. Moreover, the magnitude of the effect increased from immediate test (-17 ms, n.s.) to test on day 1 without any further training (-33 ms, $p < 0.01$). This finding is clear evidence for a lexical component in processing of pseudowords, and the development of lexical entries due to repeated incidental exposure.

Similar findings come from studies on speech perception and acquisition of words in young children (Swingley & Aslin, 2007; Gaskell & Dumay, 2003). A related approach in the field of recognition memory research was taken by Reder, Angstad, Cary, Erickson and Ayers (2002). They instructed participants to learn lists of pseudowords over a time course of 5 weeks with 2 learning session per week. Items were repeatedly presented either within one session or across several sessions. In a recognition task at the end of the sixth week, rarely presented items and frequently presented items showed a similar data pattern as low frequency and high frequency words: “low frequency” items yielded more hits and fewer false alarms than “high frequency items”. Moreover, this effect was more pronounced for pseudowords that were repeated across sessions. In sum, due to repeated presentations alone pseudowords become more and more word-like, showing lexicality effects like real words, and therefore may provide a mechanism for word acquisition and creation of a mental lexicon.

1.3 Short-term Repetition priming

Short-term priming refers to priming effects, which assumed to last only for a few seconds (cf. Tenpenny, 1995, Versace & Nevers, 2003). In short-term priming paradigms usually a task irrelevant prime is presented immediately before a target to which a response is required. Typical tasks are lexical decision, identification, word completion or categorization tasks. Priming effects were observed with identical prime and target (repetition priming, Versace & Nevers, 2003; Huber et al., 2001; Evett & Humphreys, 1981), orthographically-phonological related prime and targets (form priming e.g., De Moor, Verguts & Brysbaert, 2005; Bowers, Davis & Hanley, 2005, Versace, 1998), and semantically related prime and targets (semantic priming, e.g. Schütz, Schendzielarz, Zwitserlood & Vorberg, 2007; Kiefer, 2002, Greenwald, Draine & Abrams, 1996). In all these priming forms effects are still observable even if participants are not aware of the prime. In experimental research this can be done by masking the prime or by employing techniques like the attentional blink paradigm.

In masked priming, the prime typically is made invisible by presenting a pattern mask or letter string immediately before, after or before and after the prime. This mask renders the prime invisible despite focused attention on the prime. Nonetheless, stable priming effects were demonstrated in naming, lexical decision and identification. However, with stimulus onset asynchronies beyond 500 ms or any intervening items between prime and target these effects dissipated (Humphreys, Besner & Quinlan, 1988; Ferrand, 1996). In contrast to long-term priming experiments, masked repetition priming is not modulated by word frequency, i.e., the amount of priming is the same for low frequency words than for high frequency words (e.g. Versace, 1998, Bodner & Masson, 1997; Ferrand, 1996), and evidence for nonword priming is less clear. For example, Forster and Davis (1984) and Versace (1998) did not find stable repetition priming for nonword targets, some studies found evidence for nonword repetition priming in lexical decision (Bodner & Masson, 1997) and naming (Masson & Isaak, 1999) strongly arguing against a pure lexical mechanism and suggest that the prime helps to create a orthographic-phonological representation. More recently, Hutchison et al. (2004) compared masked repetition priming with masked form priming in a word stem completion task. Repetition priming yielded a greater effect than form priming, but regression analyses revealed that orthographic and phonologic overlap alone could account for repetition effects. Therefore, Hutchison et al. (2004) concluded that the common basis for repetition and form priming is sublexical.

Versace (1998) suggested a two component approach to repetition priming. He contrasted masked and unmasked repetition priming for words and pseudowords in a lexical decision task, and found positive priming effects for words in both priming conditions, with an interaction of word frequency and repetition only with unmasked primes. Pseudowords did not show any priming effect in masked priming and positive priming effects in non-masked priming. He proposed one short-lived and automatic component, which activates lexical entries in both masked and unmasked priming, and one component that constructs and modulates long-term memory traces, when the prime is visible. Versace and Nevers (2003) refined this approach by showing that the presentation duration of the prime is the main relevant independent variable. By manipulating the duration of an unmasked prime from 50 ms to 350 ms they replicated the findings of Versace (1998). Thus, the time consuming construction component gains weight with increasing encoding and processing time, whereas the activation component loses weight. Rendering the prime invisible by masking does not play a role.

In the attentional blink paradigm (Raymond, Shapiro & Ansell, 1992, Chun & Potter, 1995) a list of items (e.g. digits) is presented rapidly, and participants are to report two critical targets T1 and T2 (e.g. letters). Detection of T1 is quite accurate, but if the SOA between T1 and T2 are approximately ranging between 150 ms and 600 ms, detection of T2 is substantially decreased. But this decrease of awareness strongly depends on the detection of T1. If no attention is allocated to T1, detection of T2 poses no problem. Despite of this lack of awareness, T2 can induce priming effects. Shapiro, Driver, Ward and Sorensen (1997) introduced a third target letter T3 that either was the same (match trial) or different (mismatch trial) from T2. On trials, on which participants correctly reported T2, T3 was more often correctly reported on mismatch trials than on match trials, but on trials, on which participants did not report T2 correctly, T3 was more often reported on match trials. The authors concluded that if participants were aware of T2, it induced repetition blindness on T3, but if participants were not aware of T2, it induces repetition priming.

In a more recent study Visser, Merikle and DiLollo (2005) used a modified attentional blink paradigm, in which participants reported T1, made a lexical decision on T2, and afterwards performed a word completion task on the word stem of T2. On trials, on which they missed the lexical decision, they still completed the word stem with T2 above chance. This effect held for up to one second. Visser et al. (2005) concluded that unattended items are processed outside awareness. However, it cannot be ruled out, that participants consciously perceived parts of T2 (single letters or letter combination) but not the whole word, and therefore, gave a nonword response. In the word completion task the consciously perceived sublexical units could have been used to complete the word stem correctly.

Therefore short-term priming is thought to be based on a different mechanism than long term priming. Paradigms used to investigate these short lasting effects usually involve the presentation of a prime that is task irrelevant, short before presentation of the target. If prime and target word are identical facilitated responses occur. There is ample evidence that these effects are automatic in that no conscious perception is necessary. Priming effects are even observable when primes are presented subliminal.

1.4 Abstractionist vs. episodic theories on priming and word recognition

There are two broad classes of theories that can account for priming effects in word recognition and lexical decision. One class suggests *abstractionist accounts* like the logogen model (Morton, 1969), the interactive-activation model (IA; McClelland & Rumelhart, 1981, Rumelhart & McClelland, 1982), the multiple read-out model (MROM; Grainger & Jacobs, 1996) or the PDP-approach by Plaut (1997) focus on the processes that lead to word recognition. In general these models are activation based models, i.e., it is assumed that on exposure to a word its abstract representation is activated until the word is recognized. They can account for a variety of effects in word recognition, lexical decision and short-term priming. Put simply, the presentation of the prime activates its representation to a certain degree. If the target is presented shortly after the prime, residual activity in the network influences the recognition process of the target. However, it is difficult to account for the longevity and context effects of long-term priming paradigms (cf. Tenpenny, 1995, Bowers, 2000).

On the other hand, *episodic theories* like the instance theory (Logan, 1988), the counter model (Ratcliff & McKoon, 1997), as well as the Bayesian models REMI (Schooler, Shiffrin & Raaijmakers, 2001), ROUSE (Huber, Shiffrin, Lyle & Ruys, 2001) and REM-LD (Wagenmakers, Steyvers, Raaijmakers, Shiffrin, van Rijn & Zeelenberg, 2004b) are theories about implicit memory systems. They assume priming effects to be based on memory traces that are established on a first encounter, and thus stand in sharp contrast to activation based accounts. They are particularly suited to account for long-term priming and context effects, but pure episodic theories like the instance theory (Logan, 1990) have difficulties explaining inhibitory priming effects for nonwords, as well as the reversal of nonword priming effects due to variation in the speed-stress and the absence of lexicality effects in masked short-term priming. Further, episodic accounts are not capable of explaining lexical decision itself, because they are not focusing on processes. The processes on which a word-nonword decision is based are widely unknown and particularly the nonword decisions are a topic of hot debate. In the following I will shortly introduce two models that can deal with lexical decision and repetition priming, one abstractionist account and one Bayesian (episodic) account.

1.4.1 Multiple Read-Out Model (MROM, Grainger & Jacobs, 1996)

The MROM (Grainger & Jacobs, 1996) is a general model of orthographic processing claimed to account for several findings in lexical decision, naming, and perceptual identification. In essence, MROM is a derivation of the IA model (McClelland and Rumelhart, 1981; Rumelhart & McClelland, 1982) with two principles of connectionist models at its core: First, it is a parallel processing approach in that it assumes simultaneous information processing of semantic, phonological and orthographic signals, and, second, it is a competitive model in that it assumes lexical inhibition of neighbour units, i.e., simultaneous activated representations compete with each other in the identification process. According to the model, word recognition mainly operates on the lexical word form level. Therefore the semantic and phonological level can be neglected in a first approximation. The visually presented word is transformed from features to letters to sublexical units to a word form representation. According to the principles of parallel distributed processing, all similar word representations are activated to some degree and compete with each other, i.e., inhibit each other. The recognition process finishes when the activity of one word representation reaches a specified threshold. This threshold is determined solely by the stimulus and its properties, i.e., different word stimuli have different threshold settings. For example, a word that is frequent in a language has a lower threshold value than a word that occurs rarely in that language. By the time the activity of a word form representation reaches its threshold, the word is recognized correctly. If the activity of another word form representation reaches its threshold first, the stimulus is recognized incorrectly. Grainger and Jacobs (1996) assume two additional thresholds: one for the global activity, i.e., the summed activity of all word representations and a time threshold. These two thresholds are particularly important for lexical decision: For a correct lexical decision, complete recognition of a word is not necessary. Rather, it is sufficient to consider the global activity. If this activity reaches its threshold, a ‘word’-response is generated, if the temporal deadline is reached before either the global activity threshold or the word representation threshold is reached, a ‘nonword’-response is generated. All three thresholds – the ones for single word representation, global activity and the temporal deadline – are noisy in that their values vary around a mean from trial to trial. The global activity threshold and the temporal threshold are further adjustable by strategic effects, response criteria, and experimental context.

MROM has received support from various studies in the last decade (e.g. de Moor, Verguts & Brysbaert, 2005) and has been developed further to account for phonological effects (Jacobs, Rey, Ziegler & Grainger, 1998). However it has not been applied to repetition priming yet. De Moor et al. (2005) applied MROM to masked priming: On presentation of a prime its representation is activated and inhibits competing representation of similar words. This activation and inhibition still last, when the target is presented, and so influences its processing. In the same way repetition priming should influence the time it takes to reach one of the three criteria. For example, when a word is presented the first time, its representation is activated and inhibits neighbour representations. On the next encounter of the same word it takes less time to reach a threshold, which leads to a faster word response. But what does MROM predict for nonword trials? MROM assumes that a high global activity gives rise to a longer temporal deadline (Grainger & Jacobs, 1996). Further it is assumed that (a) on nonword presentation several similar word representations are activated to some extent and (b) that these are still active on the second encounter of the nonword, so that the global activity is increased compared to the first encounter, and the temporal deadline is extended. Thus, longer response times for correct identified nonwords are to be expected.

1.4.2 Retrieving effectively from memory – lexical decision (REM-LD)

The REM-LD (Wagenmakers et al., 2004b) model is a derivation from the general memory model REM (e.g. Shiffrin & Steyvers, 1997) and highly similar to REMI (Schooler et al., 2001), a model for implicit memory in word identification. At its core, REM-LD proposes a Bayesian decision process that weighs the word evidence against the nonword evidence. From the REM model it adopts the assumption that, on exposure to a known word, (a) a specific episodic trace of the encounter is stored in episodic memory and (b) simultaneously the existing lexical/semantic trace of the specific stimulus is updated with information about the stimulus content and context. Crucially, only information that is not already part of the semantic/lexical trace is added. On a lexical decision trial the decision probe is matched simultaneously to several traces from long-tern semantic lexical memory on a feature-by-feature base. The decision probe consists of the presented stimulus and its present context. However, it is not identical to the physical stimulus, but rather is a noisy representation of this stimulus. Therefore, probe and lexical/semantic trace do not have to match, even if they encode the same item. REM-LD computes (a) the probability that decision probe and a lexical trace match given they represent the same item, and (b) the probability that decision probe and a lexical trace match given they represent different items.

In lexical decision a word item always corresponds to exactly one lexical/semantic trace. In contrast, a nonword item never corresponds to any lexical/semantic trace. An activation function determines how many features of a probe are active at time t, determining the actual probabilities. From these probabilities the posterior odds ratios that the probe is a word can be computed, on which the word vs. nonword decision is based. Applied to repetition priming REM-LD can be described as follows: On the first encounter of a word its lexical/semantic trace is retrieved form long-term memory, and information about the stimulus and context is added. On the second encounter of the same word, the lexical/semantic trace again is retrieved from long-term memory, this time with more information that matches the stimulus probe, becoming evident in a greater probability that decision probe and lexical/semantic trace match. A contrary process is assumed for nonwords: Exposure to a nonword leads to activation of several similar word traces. To one of these traces, information is added in the lexical trace, increasing the probability that this trace and the decision probe of the nonword match, although encoding different items. So, this mechanisms of REM-LD predicts inhibitory repetition priming for nonwords.

Two aspects of REM-LD are important for the present purpose: First, because just new information is added to a lexical trace on word presentation, multiple presentations of the same word lead to increasing effects, that eventually reach an asymptote, i.e., each presentation contribute less to the effect than the preceding presentation. Second, because part of the new information added to the lexical trace is assumed to be context, multiple word presentations lead to larger effects when the contexts are most distinct, e.g. if two presentations are presented successively with only two items between, the contexts are almost the same. This means that on the second presentation nearly no new information is added to the trace. In contrast, if two presentations are intervened by 20 items the contexts differ more, and thus, the second presentation adds more new information to the lexical trace. In this regard REM-LD predicts different results than MROM, because an activation approach predicts decreasing priming effects with increasing lag between two prime presentations.

1.5 Neuronal correlates of repetition priming

Since behavioural evidence points to several different forms of repetition effects the question arises to what degree this approach is supported by neurophysiological evidence. In general, repetition of a stimulus is accompanied by a reduction of neuronal activity, as confirmed by studies in single cell recordings, functional magnetic resonance imaging (fMRI), and electroencephalogram recordings (EEG) in primates and humans (for a review see Grill-Spector, Henson & Martin, 2006).

In ERP research on repetition priming, the amplitude of the waveform for a repeated stimulus is more positive than for a stimulus on its first occurrence. Effects can be observed in a time window of ~150 ms to ~600 ms. The earliest components occur focally at posterior right-hemisphere scalp sites, and are supposed to reflect early sublexical orthographic processing (Holcomb & Grainger, 2006). Most common is a more positive waveform for repeated stimuli at 400 ms (N400) after onset (Holcomb & Grainger, 2006, Friedrich, 2005, Karayanidis, Andrews, Ward & McConaghy, 1991), which is also a prominent marker of semantic priming (e.g. Kiefer, 2002).

Deacon, Dynowska, Ritter and Grose-Fifer (2004) have shown N400 repetition effects for pronounceable nonwords that were neighbours of real words, as well as for pronounceable nonwords without any word neighbours. However, these studies did not distinguish between different forms of repetition priming. Guillem, Rougier and Claverie (1999) recorded intracranial ERPs from several brain structures, while patients performed an object recognition task, in which stimuli were repeated at two different lags (6 vs. 19 intervening items). Structures in the posterior cortical areas, the lateral frontal lobe and the anterior cingulate gyrus showed reliable N400 effects largest in the short lag condition, whereas orbitofrontal areas, hippocampus and amygdala showed N400 effects that were largest in the long lag condition, and anterior temporal structures showed effects in both short and long lag conditions. The authors take this as evidence for a short-term semantic network and a long-term episodic memory network. Henson, Rylands, Ross, Vuilleumier and Rugg (2004) demonstrated an early (150ms – 300ms) and a late component (400 – 600 ms) in repetition priming of objects. With increasing lag the magnitude particularly for the early component disappeared. However, differences were only quantitatively, in that the magnitude for both components was reduced, revealing no dissociation between different priming mechanisms for different lags.

In fMRI, repetition effects manifest as reduction in the BOLD signal at locations widely distributed over the cortex but pronounced in the temporal-occipital region, the fusiform gyrus, and the visual word form area (VWFA, see Grill-Spector et al. 2006; Brozinsky, Yonelinas, Kroll & Ranganath, 2005; McCandliss, Cohen and Dehaene, 2003). This reduction in BOLD signal as a neural marker for stimulus repetition has been demonstrated in object naming (van Tourenout et al., 2000, Vuillemier, Schwartz, Duhoux, Dolan & Driver, 2005), visual word recognition (Dehaene, Naccache, Cohen, et al., 2001) and lexical decision on spoken words and pseudowords (Orfanidou, Marslen-Wilson & Davis, 2006).

Henson, Shallice and Dolan (2000) employed a continuous search task in which participants watched a stream of familiar and unfamiliar faces or symbols; they had to press a key whenever an inverted face or an exclamation mark occurred. Items of interests were non-response stimuli, which were repeated at varying lags. On the second presentation of familiar objects, activity in the right fusiform area was reduced as compared to the first presentation; for unfamiliar objects, however, activity was enhanced. This suggests qualitative differences in the processing of repeated familiar and unfamiliar objects, and may hint at different mechanisms for word and nonword repetition as well. Henson et al. (2000) further demonstrated that these activity patterns continue over several repetitions. The BOLD signal for familiar objects steadily decreased with increasing presentation frequency but steadily increased for unfamiliar objects. In contrast, in a similar study, Reber, Gitelmann, Parrish and

Mesulam (2005) did not find a BOLD reduction beyond the second presentation. However, both studies lack from behavioural data, and so no direct relationship could be established from neural priming to behavioural priming.

Dobbins, Schnyer, Verfaellie and Schacter (2004) investigated the neural basis of long-term repetition priming in more detail, and disentangled response learning and repetition components. They had participants categorize objects as ‘bigger than a shoebox’ or ‘smaller than a shoebox’. Objects were either repeated or not. After the first phase, the categorization rule was reversed, so that a repeated stimulus now required a response contrary to that on the first encounter. In the third phase, the rule returned to the starting rule. In the first phase, behavioural priming was observed as well as a BOLD reduction in the fusiform area and prefrontal cortex (PFC). Switching the rule reduced behavioural priming and neural priming in the PFC, but completely eliminated neural priming in the fusiform area. Further, the reduction of neural priming in the PFC correlated with the reduction of behavioural priming. These findings highlight the role of response learning in repetition priming. Especially in long-term priming paradigms (e.g. continuous lexical decision tasks) the same response is required on multiple repetitions of the same stimulus. Thus, a great deal of repetition priming effects can be explained by response learning. Wig, Grafton, Demos and Kelley (2005) confirmed the relationship between activity in the PFC and behavioural long-term priming. Transcranial magnetic stimulation (TMS) of the left inferior frontal cortex during the first and second presentation of an object abolished behavioural long-term priming, as well as neural priming in the PFC and posterior temporal areas but not in occipital cortex.

Taken together, neurophysiological evidence supports the notion of dissociable forms of repetition priming regarding long-term (episodic) effects and short-term (semantic/lexical) effects, as well as dissociable mechanisms for familiar to unfamiliar items. Further, the importance of response learning is highlighted, which has to be considered in the interpretation of long-term priming observed in paradigms that require the same response on repeated items.

1.6 Open questions

Most studies on repetition priming investigated effects on the first repetition. Studies employing more than one repetition have mostly been conducted in long-term priming research. However, the question remains to what extent priming is modulated when multiple repetitions of words are presented in a short-term priming paradigm. Abstractionist accounts as well as episodic accounts predict a cumulation of priming effects. So is it possible, to induce long-lasting priming effects by extremely short but multiply repeated presentation of words?

A second question concerns the vulnerability to interference: Can priming still be observed when items are presented for extremely short durations with several other intervening items impeding the processing of each single prime? In long-term paradigms there is evidence for a stable repetition component and a component that decays with more than 2-3 intervening items (McKone, 1995, 1998; Ratcliff et al., 1985). On the other hand, studies of processing of items presented in a rapid serial visual stream have shown that priming in a stream is possible, but decays very fast (Visser et al., 2005). Studies that obtained longer lasting effects employed very short lists (6 – 10 items) with a relative high presentation duration (~200 ms – 1000 ms) and critical items placed at specified positions near the end of the stream (Masson & MacLeod, 2000; Masson & Hicks, 1999).

A third question of main interest concerns nonword repetition effects: In long-term priming nonword targets yield facilitatory and inhibitory priming results, depending on the type of study task, whereas short-term priming paradigms yielded mixed effects for masked priming and facilitatory effects for unmasked priming. What effect do multiple repetitions of non-target nonwords have on subsequent lexical decisions? In light of results from word acquisition research, nonwords should become more and more word-like and, eventually, should be represented on a lexical level. Therefore, inhibitory priming effects should occur.

A fourth question concerns the importance of response selection. Repetition effects for items that had been selected for response once were shown to last for days and even longer, so the question remains if such effects can be obtained from repeated nontargets as well.

1.7 The account of this study

The experiments presented in the following chapters were designed to answer the above outlined questions, and to develop a priming paradigm, which is capable of inducing priming effects in several items in one sweep. For illustration purposes one may start with a simplified informal MROM model. According to the model, the presentation of a word (e.g. “REGEN”) leads to the activation of its abstract representation on the lexical level. Simultaneously the lexical representations of orthographically similar words are activated, but to a lesser degree (e.g. “SEGEN”, “REGEL”,”ROGEN”). The same happens when an orthographically related pseudoword is presented (e.g. “RELEN”). In this case, the representations of “SEGEN”,”REGEL”,”ROGEN”, and “REGEN“ are activated to a small amount. However, the summed activation of all word representations is less than on presentation of “REGEN”, because none of the representations completely match the pseudoword stimulus.

When only one word is presented, the activation of its representation deceases over time (Figure 1, upper panel), so that at time t essentially no activity remains. By presenting the same word multiple times the activation cumulates and becomes persistent, making the lexical representation more immune against competitive representations, so that at time t the word representation is in an active state still (Figure 1, middle and bottom panel). In this case a lexical decision on the multiply presented word should be facilitated. So it might be possible to induce long and stable priming effects by presenting repetitions of the same word in rapid succession. According to the model, two word representations that are active simultaneously inhibit each other, so that only the “winner unit” is persistent over time. However, it is reasonable to assume that inhibition is stronger between representations of orthographically similar words (e.g. “REGEN” – “SEGEN”) than for representations of orthographically dissimilar words (“REGEN” – “FLUCH”). Thus, it should be possible to induce persistent activity in several words simultaneously.

The experiments presented here differ in several ways from previous investigations of repetition priming. First, I employed a prime presentation of short duration on the first encounter as used in short-term priming, but I test effects in the time range of up to minutes as done in long-term priming studies. Second, priming studies using rapid serial presentations always used extreme short lists with two presentations of one critical item at the end of the lists. In my study, the lists are considerably longer in duration, with many more items and several critical items within one stream. Thus, the purpose of my study was to develop a priming paradigm that combines methodological aspects of short-term paradigms, long-term paradigms, and rapid serial presentation outlined above. From long term paradigms, the distinction of a study phase and a test phase and the long time span between study and test was adopted. But prime presentation in the study phase was more similar to short term paradigms: primes were presented for very short durations and did not require a response, and processing of primes was degraded by the high presentation rate. I presented participants rapid word streams of more than 200 items rushing by within less than a minute. According to the ideas outlined above, an increasingly facilitatory positive repetition priming effect should be observed for words. This would also be in line with previous results that have shown cumulative priming effects in long-term priming (Wagenmakers et al. 2004). However, multiple presentations of pseudowords in a rapid stream should lead to increasingly inhibitory priming effects.

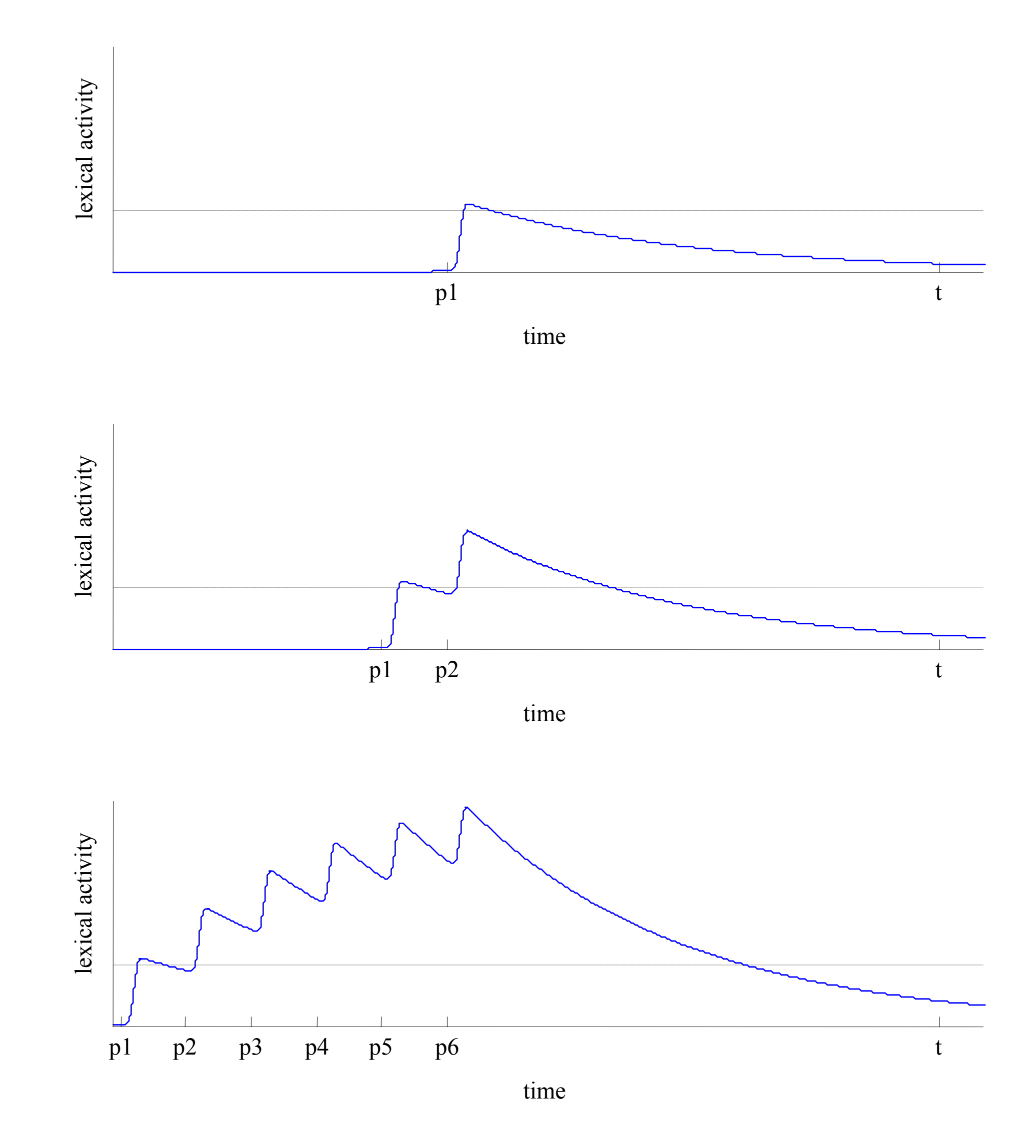

Figure 1: The assumptions underlying the present experiments. Presentation of a word stimulus (p) leads to activation of its lexical representation, which deceases over time so that at time t only little activity remains (upper panel). Multiple presentations of the same stimulus lead to cumulative lexical activation, so that at time t more residual activity remains, although the time delay from the last presentation to time t is always the same (middle and bottom row).

In Experiments 1a and 1b the basic approach is introduced in detail. It is investigated whether the multiple presentation of a word or pseudoword in a rapid stream leads to repetition priming effects on a later lexical decision on the same word or pseudoword, even if the presentation in the stream is not response-relevant. The construction of word streams is evaluated, and the influence of presentation duration of primes is investigated. Experiment 2 tries to disentangle priming influences from recognition components by manipulating the probability of awareness of single primes. Experiment 3 was conducted to test the influence of the level of processing of repeated words and pseudowords during the word stream. In addition, recognition performance of primes is quantified as a function of presentation frequency.

In Experiment 4 the longevity of priming is investigated, by contrasting an immediate test phase following the word stream with a delayed test phase with another word stream intervening. If priming arises from transient activation, the amount of priming should be less with delayed testing. Experiment 5 compares repetition priming with cross-priming of words on pseudowords and vice versa. If pseudowords lead to activation of representation of orthographically similar words, then cross-priming should be observed.

Experiment 6a and 6b test a variant of the massive repetition paradigm, in which short word streams with one critical word only are used. This simplifies the control of interfering variables because distractor items can be chosen more carefully. Furthermore, a very long lag of 24 hours is introduced. Here, repetition priming for items that had been targets on the preceding day, and items that had been distractors on the preceding day are compared As argued in Chapter 1.5, response learning may play a crucial role in long term repetition priming, so it is questionable if non-response primes can also induce long lasting effects.

The questions can be summarized as follows:

1. Can repetition priming be induced for words and pseudowords simultaneously by presenting a word stream of 200 items within one minute? Does priming cumulate beyond the second presentation? Does massive repetition priming depend on awareness of a singular presentation?

→ These questions are investigated in Experiments 1 to 3

2. How long does the massive repetition effect last? Is there a qualitative difference between very long lag priming for response selective primes and non-response selective primes?

→ This question is investigated in Experiments 4 and 6a

3. Does processing of pseudowords activate the word form representations of neighbour words?

→ This question is investigated in Experiment 5

4. Do short word streams with one critical item only and with immediate test enhance the effects?

→ This question is investigated in Experiment 6b.

In sum, the present experiments try to establish a technique for priming research that is capable of inducing priming effects in a large body of items simultaneously. The possibilities and limitations of this technique will be investigated. Results obtained with this technique are relevant regarding the discussion about different components of repetition priming and the processing of pseudowords.

1.7.1 Data analysis and the "language-as-fixed-effects-fallacy"

One often neglected problem in analysis of psychological research is the so called 'language as fixed effect fallacy' (Clark, 1973). Clark (1973) pointed out that in psycholinguistic research analysis with just Subject as random factor is insufficient, because not only participants are randomly chosen, but items, too. In general, experiments are designed with few participants in repeated measurement, i.e., each participant is measured repeatedly under several experimental conditions with randomly chosen items. For analysis, data for each participant are collapsed over items and entered into a repeated measurement ANOVA with Subject as random factor. In such cases, it is neglected that items (i.e., words) were sampled from a population, too. This random selection and allocation to different experimental conditions leads to sample variance that is neglected in conventional repeated measurement ANOVA, differences between items are simply ignored. Beside the problem of unjustified generalization, a more profound problem is inflation of Type I error (Clark, 1973; Raaijmakers, Schrijnemakers & Gremmen, 1999, Raaijmakers, 2003). In several simulation studies, Rouder and Lu (2005) showed that the real Type I error rate increases with item variability. Moreover, the inflation effect is even more profound when the sample sizes of items and participants are increased. To account for these problems, it is common to conduct a subject analysis (F1) and an item analysis (F2), and to reject the null hypothesis if both F1 and F2 are significant. But this procedure neither makes generalization more feasible nor does it eliminate the alpha inflation problem, because both F1 and F2 are biased (Raaijmakers et al., 1999, Raaijmakers, 2003). Several solutions had been developed in the last decades including a quasi F-ratio statistic (Clark, 1973), counterbalancing of items (Raaijmakers et al. 1999, Raaijmakers, 2003), hierarchical Bayesian approaches (Rouder & Lu, 2005). Two more recent approaches are the so called linear-mixed-effect models (LME models), which can be subdivided in hierarchical multi-level-analyses (Quené & van den Bergh, in press) and LME models with crossed random factors (Pinheiro & Bates, 2000; Baayen, Tweedie & Schreuder, 2002; Baayen, Davidson, & Bates, submitted).

The great advantage of LME models is that they can deal with several "crossed" random effects and fixed effects as well as with nested random factors, i.e., Subject and Item not only vary in intercepts but also in slope. Thus, by modelling the data with LME models F1 and F2 analyses can be conducted in one sweep. For a detailed description of all computational and theoretical aspects of linear mixed effect models see Venables and Ripley (2002) and Pinheiro and Bates (2000). Beside the possibility of crossed random factors LME model have other advantages compared to conventional repeated measures ANOVA. First, no assumptions about sphericity and compound symmetry are necessary, because the variance covariance matrix can be modelled directly from the data; second, the power of LME models has been shown to be superior to both conventional repeated measures ANOVA and multivariate repeated measures ANOVA, even in cases of heteroscedasticity; and third, LME models are robust against missing data (Quené & van den Bergh, 2004). A problem might arise from the fact that (1) response time (RT) data are not distributed normally, but rather show high right-skewness and (2) that the variance of RT usually increases with increasing response times. This poses no problem in the usual case of repeated measures ANOVA, because individual reaction times are aggregated for each participant. Therefore, *mean* reaction times for participants are entered into the analysis, which leads to a decrease in the skewness of distributions. In mixed effects LME models with individual measurements entered into the analysis, it remains unclear what impact the violation of normal distribution of residuals has. In such cases it is common to use an appropriate transformation of RT data (e.g. reciprocal RT). However, it is highly questionable whether the results of such an analysis can be interpreted in terms of response times because the inverse transformation of response times changes the scale and is hardly neutral regarding the content (e.g. equal differences are not equal when transformed inversely). Therefore, a significant interaction of raw RT need not to be significant when analysed with reciprocal RT. Thus, it must be checked whether such transformations alter the interpretation of results. Figure 2 illustrates this for real RT data (left panel) and their reciprocals (right panel). Homoscedasticity as well as a normal distribution of residuals are far better met by reciprocal RT. Because of the problems of interpreting analyses of inverse RT in terms of absolute RT mentioned above, I checked to what extent a transformation changes the data pattern. Individual response times were transformed, then the mean and standard error were computed, and these were then re-transformed into absolute RT values. By comparing plots of the original RT and re-transformed RT the amount of distortion becomes visible. Figure 3 shows data from Experiments 1a and 1b in untransformed, inversely transformed and re-transformed conditions. Considering the inverse re-transformation data the RT range is slightly compressed, but the data pattern as such is clearly replicated.

Another minor drawback might be that the appropriate number of degrees of freedom for the fixed effects in a LME model is unknown. Only an upper bound can be reported, which does not take the additional random factors into account. So no p-values or confidence intervals can be reported (Baayen et al., submitted). However Baayen et al. (submitted) suggest that for large data sets a "quick and dirty" analysis is to check that the t-values exceed 2. A more convenient method is Markov chain Monte Carlo sampling from the posterior distribution parameters. From this sample the highest probability density (HPD) confidence intervals can be computed for fixed effects (Baayen et al., submitted). For random effects, the authors suggest likelihood ratio tests for comparison of two models, one with the random factor in question and one without. The general approach chosen in the presented experiments largely follows the recommendations of Baayen et al. (submitted) for use of LME models with crossed random factors. A detailed outline for analysis of RT data is described in the next section, deviations from this outline and specific aspects of data analysis will be added in the description of the relevant experiment.

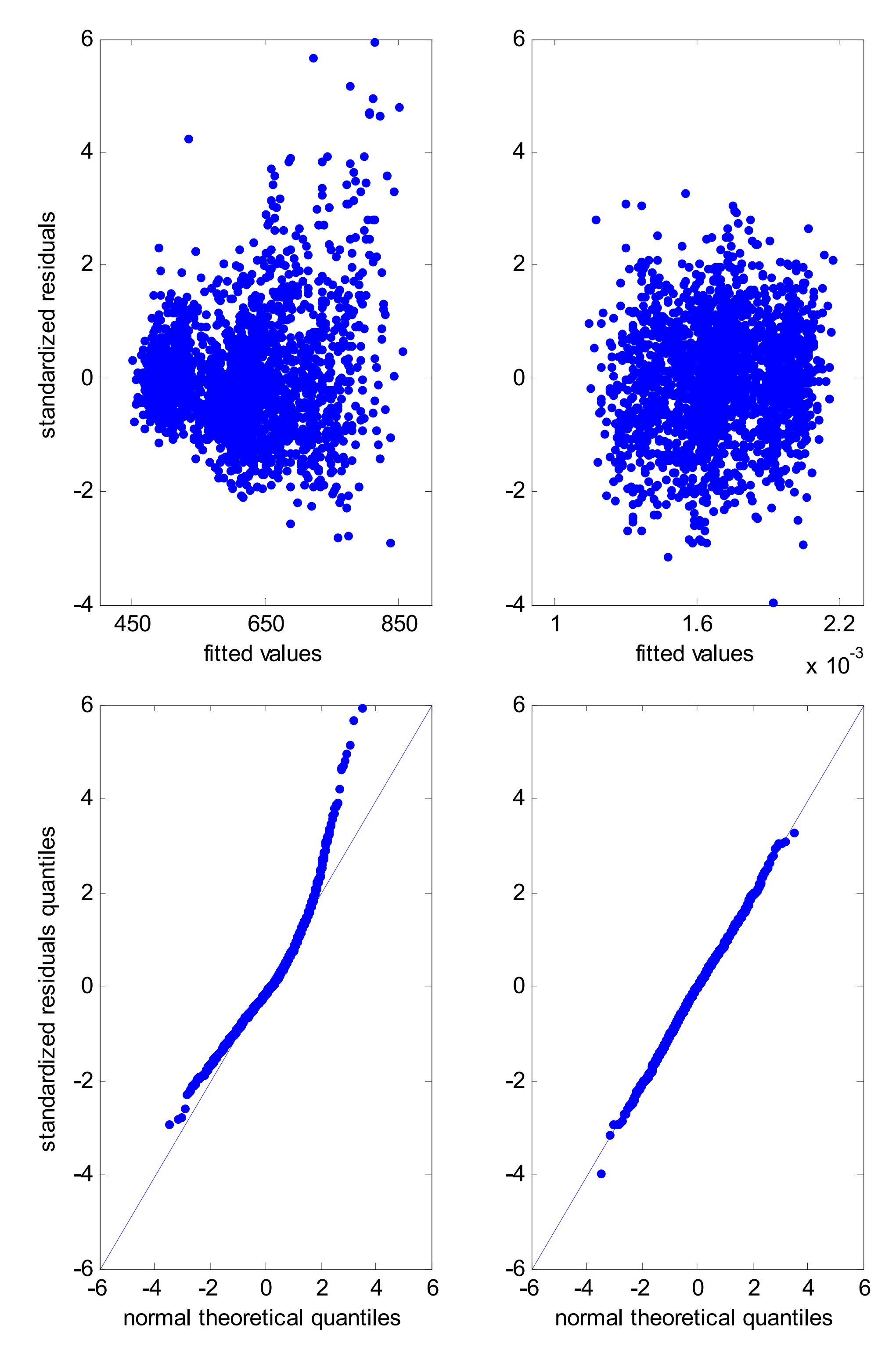

Figure 2: standardized residuals plotted against fitted values (upper row) and Q-Q-plots (bottom row) for Experiment 1a: RT (left column) and inverse RT (right column).

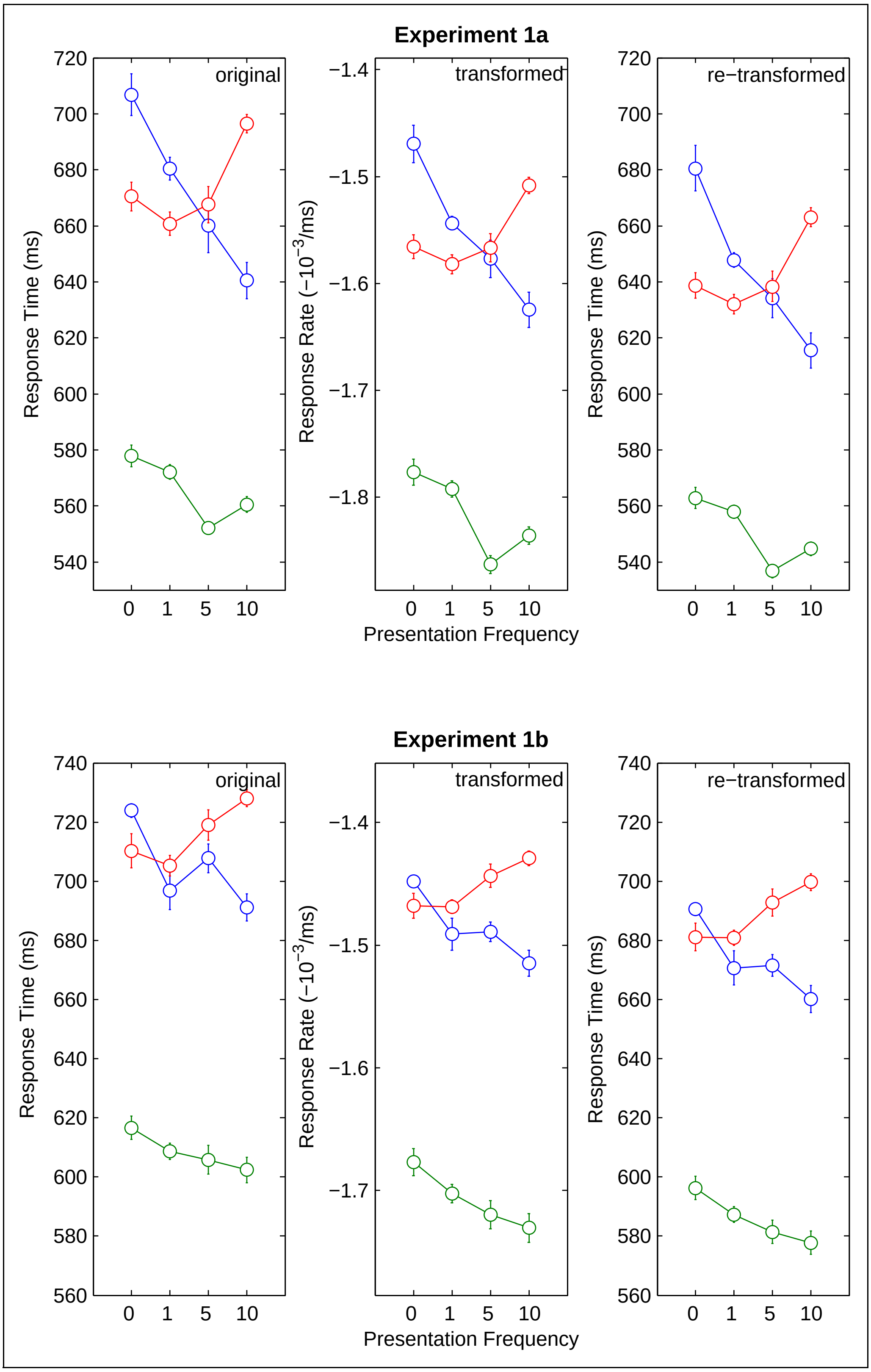

Figure 3: Exemplary data from Experiment 1a (left column) and experiment1b (right column). The comparison of original data (upper row), inverse transformed data (middle row) and re-transformed data (bottom row) shows that RT range is compressed, but the data pattern is replicated.

1.8 General Method

Unless mentioned otherwise, all experiments are based on the methods and materials described in the following.

Task. The experiments consisted of two consecutive tasks: a *counting task* and a *lexical decision task*. In the counting task, participants had to focus attention on a stream of rapidly presented words and pseudowords and to count the number of forenames occurring in the stream. After each stream, participants performed several lexical decision trials on single words and pseudowords, some of which had been presented within the preceding word stream, and some were new words and pseudowords.

Stimuli. Low frequency (LF) and high frequency (HF) word stimuli were selected from the WebCelex database (Max Planck Institute for Psycholinguistics, 2001). All were German nouns. LF and HF words were matched for length; all were 3-6 letters long. Names of people, places, etc. were discarded.

Pseudowords were created by randomly replacing one letter of an existing German noun, obtained from WebCelex (Max Planck institute for Psycholinguistics, 2001), e.g. the word WURZEL becomes WURWEL. To ensure the legality of each pseudoword two constraints were included. First, the consonant-vowel-structure of each word was maintained, i.e., vowels were replaced by vowels and consonants by consonants, so that each pseudoword was pronounceable. Second, as measure for the “word-likeness” of pseudowords, a trigram frequency distribution was computed from 51.728 entries in the German lemma section of WebCelex (Max Planck institute for Psycholinguistics, 2001). Each trigram in a pseudoword had to occur at least once in this distribution. For each pseudoword, the mean trigram frequency was taken as “frequency count”. Further, from each parent word 20 pseudowords were calculated, and only the one with the highest trigram frequency count was selected. More information on stimulus parameters is given in the description of each experiment.

The counting set comprised 150 male and 150 female forenames.

Design. An experimental session was subdivided into several runs. One run consisted of a single word stream and a block of several lexical decision trials. Each word stream comprised an equal number of words and pseudowords. Half of the words were LF-words and the other half were HF-words. Repetitions of items were classified in (a) *repetitions within* word streams and (b) *repetitions across* word streams. Items repeated within word streams either occurred 1, 5 or 10 times within one word stream. Items repeated across word streams occurred once in every word stream of a session. Figure 4 illustrates the design of word streams: In the upper panel the word “MESSER” (knife) occurs repeatedly within one word stream, the pseudoword “WURWEL” only once. Both words are tested in the following block of lexical decisions. In the lower panel the word “FREVEL” (sacrilege) occurs once in every stream, the pseudoword “BLABE” occurs once in the last stream. Both are tested in the lexical decision block that follows the last stream. The mean spacing between two repetitions of the same item (and between two once-presented items) was 13 items, with a random jitter of plus/minus 2 items (11 items +- 2 in Experiment 1a and 1b). Word streams were constructed so that the mean of the mean stream position for each presentation frequency was roughly equal. In addition, 10 counting items were randomly chosen (without replacement) and placed in random positions. The spacing between two consecutive names followed a triangular distribution with a mean of 21 items. Filler items were assigned to the first and last two positions of every stream to ensure that none of the critical stimuli were presented in the very beginning or ending of a stream.

The number of lexical decision trials per block varied from experiment to experiment. Targets were words or pseudowords that had occurred in the preceding word stream, and an equal number of words and pseudowords that had not been presented before.

Word Type and Presentation Frequency (PF) within and across word streams were the independent variables of main interest. Words and Pseudowords were treated separately, and were not compared. The main dependent variable was RT in lexical decision. Error rates were not considered as dependent variable because lexical decisions were performed under low speed stress, which led to extremely low overall error rates for high frequency words and pseudowords.

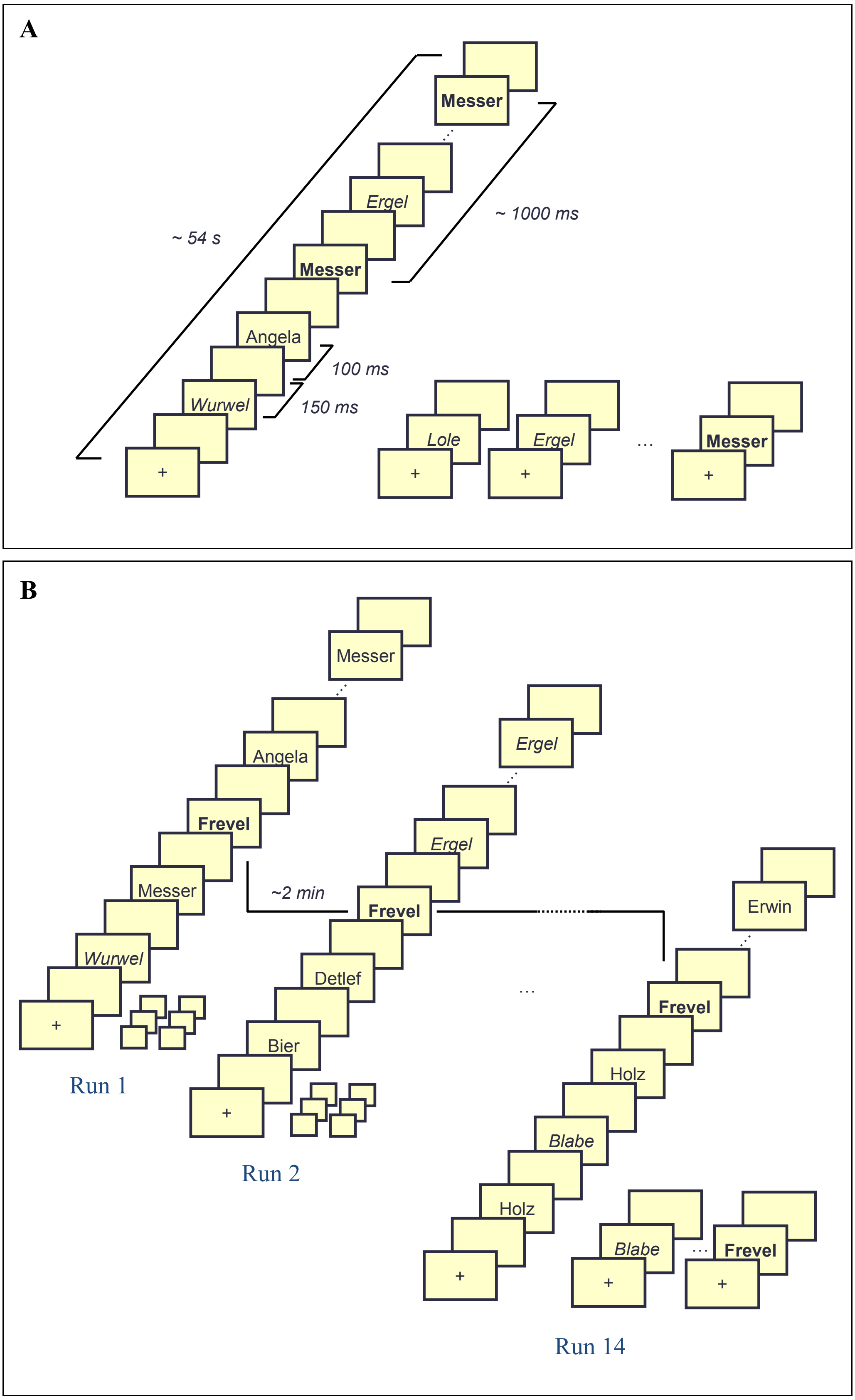

Figure 4: Construction of word streams. A: Exemplary illustration of one word stream with repeated presentations of the word "Messer" (left) and the following block of lexical decisions (right). B: Sequence of several word streams with repeated presentations across word streams of the word "Frevel". For illustration purposes the critical words are printed in bold and pseudowords are printed in italic. In the experimental setting all items were presented in standard font.

Apparatus. The experiments were run on a Microsoft Windows based PC connected to a 17” Iiyama monitor with a resolution of 1024 X 768 pixel and a refresh rate of 75 Hz. Stimuli were presented in black on a grey background (RGB values 209, 209, 209). Font was “Verdana”, font size was set to 30 points. Participants were seated in a distance of 65 cm, with the head arrested by a chin rest. Stimulus size ranged from 0.84° x 1.34° visual angle (three letter words) to 0.84° x 4.61° visual angle (six letter words). Stimuli order and presentation were controlled by MATLAB 7.1 and Presentation 9.9, respectively.

Data preprocessing. Incorrect trials were excluded from further data analysis. To ensure that analyses were not greatly influenced by outliers, trials with reaction times that deviated more than 2.5 standard deviation units from the individual mean reaction time for each condition were discarded. Response times were transformed to their reciprocal value (i.e., response rate). For the reasons of data preprocessing see Chapter 1.7.1. Unless mentioned otherwise, results are reported in terms of re-transformed response times: Individual response times were transformed to their reciprocal value. The mean reciprocal response time was then re-transformed to response time scale.

Model fitting. Data analysis was conducted with the function *lmer* included in the *lme4* package (Bates, 2007) and several other functions from the *languageR* package (Baayen, 2007) both supplied in the R system for statistical computing (version 2.4.1, R Development Core Team, 2007). For words and pseudowords, separate LME model were fitted to the reciprocal response times with crossed random factors for Subject and Item. Fixed factors and predictor variables varied from experiment to experiment, but in general Word Type/Pseudoword Type and Presentation Frequency were included as factors, and Trial Number was included as continuous predictor variable. Linear contrasts were tested for the ordered factor Presentation Frequency; for all other factors difference contrasts were tested. Because of the difficulty in determining appropriate degrees of freedom (see Chapter 1.5.1), *t*-values are reported without them; *p*-values were computed using Markov Chain Monte Carlo sampling (n = 10000 simulations) for the fixed factors. In addition, likelihood ratio tests were conducted in each case to test the contribution to the model fit of the random factors and of continuous predictor variables like trial number.

2. Experiment 1a and 1b: Basics and details of the method

2.1 Experiment 1a

The aim of Experiment 1a was to demonstrate how repetition priming accumulates when words and pseudowords are massively presented in a word stream and tested once in a following lexical decision task. Is there cumulative priming beyond the second presentation? Can more than one item be primed simultaneously? Cumulative priming has been shown by Wagenmakers et al. (2004) in a long-term paradigm, where each word was target in lexical decision for five times. Masson and Hicks (1999) showed with short lists of six items that non-target words in a list of rapidly presented items produced priming effects in a later lexical decision task when items were presented for 200 ms. In Experiment 1a the presentation duration of single items in the word stream followed that of Masson and Hicks (1999), but the stream length and duration was extended to 214 items and approximately one minute. I hypothesized that priming facilitates the 'word'-response for words, but impedes the 'nonword'-response for pseudowords, and that these effects increase with increasing presentation frequency within a word stream. Further I investigated if there still is reliable repetition priming, when an item is repeated across word streams, so that the lag between repetitions increases to several minutes.

2.1.1 Methods and Materials

Participants. Six students (one male; age 21 - 25 years, mean 22.7 years) from Technical University Braunschweig were tested in two one-hour sessions at least two days apart. All were native German speakers and took part for course credit. Vision was normal or corrected to normal.

Task. Participants performed the *counting task* and *lexical decision task* as described in the General Method section.

Stimuli. 550 LF-words, 550 HF-words, and 1100 pseudowords were used. LF-words and HF-words had frequency counts of less than 2 per million, or more than 16 per million, respectively (for more details see Table 1 and 2). For each participant and each condition, words and pseudowords were randomly chosen from the target set without replacement.

Table 1: Stimulus characteristics in Experiment 1a and 1b.

	Letters *mean (SD)*	**Syllables** *mean (SD)*	**CELEX frequency** *count/million*	*log/million*	**Trigram frequency**
LF-words	5.0 (0.9)	1.7 (0.5)	0.8 (0.5)	0.08 (0.22)	158.4 (18.3)
HF-words	5.0 (0.9)	1.6 (0.5)	97.2 (159.1)	1.75 (0.40)	168.2 (14.0)
Pseudowords	5.0 (0.9)	1.7 (0.5)	--	--	161.4 (14.7)

Table 2: Mean values (SD) of letter overlap within and between stimulus classes in Experiment 1a and 1b.

	LF-words	**HF-words**	**Pseudowords**
LF-words	0.28 (0.19)	0.28 (0.20)	0.30 (0.20)
HF-words		0.29 (0.20)	0.30 (0.20)
Pseudowords			0.32 (0.20)

Design. Each word stream contained 214 items. There were 48 words, 48 pseudowords, 10 forenames and 4 filler items (2 words and 2 pseudowords), which sum up to 110 different items per stream. Table 3 shows the number of occurrences of each item type per stream. Note that (a) half of the once presented items were the same words in every stream, and that (b) in run 1 to 13 only 2 LF-words, 2 HF-words and 4 pseudowords were tested.

The experiment consisted of 2 sessions which were identical in design and procedure. A session consisted of 14 runs. In the first 13 runs the lexical decision block comprised 32 trials, in which only words and pseudowords were targets which had been presented either in the preceding word stream or not presented at all. In the 14^{th} run all words and pseudowords that were presented once in each word stream, together with an equal number of words and pseudowords presented once in the last word stream were tested. In sum, the last lexical decision block consisted of 80 items.

Table 3: Number of occurrences of each item type per word stream

	Number of presentations			
	1*	**5**	**10**	**Σ**
LF-words	20	2	2	24
HF-words	20	2	2	24
Pseudowords	40	4	4	48
Number of different words	80	8	8	96
Number of presentations	80	40	80	200
Names				10
Filler items				4
Σ				214

* Half of the items were the same across all streams. In runs 1-13 only 1/10 (2/2/4) items was tested, in run 14 all items were tested.

Procedure. Each stream started with a fixation cross presented for 100 ms centred on the screen, followed by a blank screen presented for 500 ms. Then the items of the stream were presented in rapid serial order. All items were centred on the screen. Each item was presented for 150 ms, and was followed by a 100 ms blank. Thus, the whole stream lasted for approximately 54 seconds. Immediately after, participants reported the number of names counted during the stream presentation via the keyboard.

After the number report, participants started the lexical decision block by pressing a button. Lexical decision blocks consisted of 32 decisions in Run 1 to Run 13, and 80 decisions in Run 14. Each trial began with a fixation cross for 1000 ms, followed by a blank that lasted for 500 ms. Then the target word or pseudoword appeared for at most 1500 ms or until the response. Slow or incorrect responses were signalled visually by “ZU LANGSAM!” and “FEHLER!”, respectively. After the response or the feedback, a blank screen was shown for 1000 ms, before the next trial began with the fixation cross. At the end of a lexical decision block, a summary feedback was provided with the relative amount of incorrect and slow responses. If the amount of incorrect or slow answers exceeded 5 %, participants were reminded to be faster or more accurate.

2.1.2 Results

Counting task. Mean number of name counts was m = 8.7 (median = 9), i.e., on average participants missed 1 to 2 names per word stream. However, there were considerable differences between individuals: Four individuals achieved nearly perfect results (9.3 < m < 10.1), and two participants deviated considerably from the actual number of 10 items (m = 7.9 and m = 5.9).

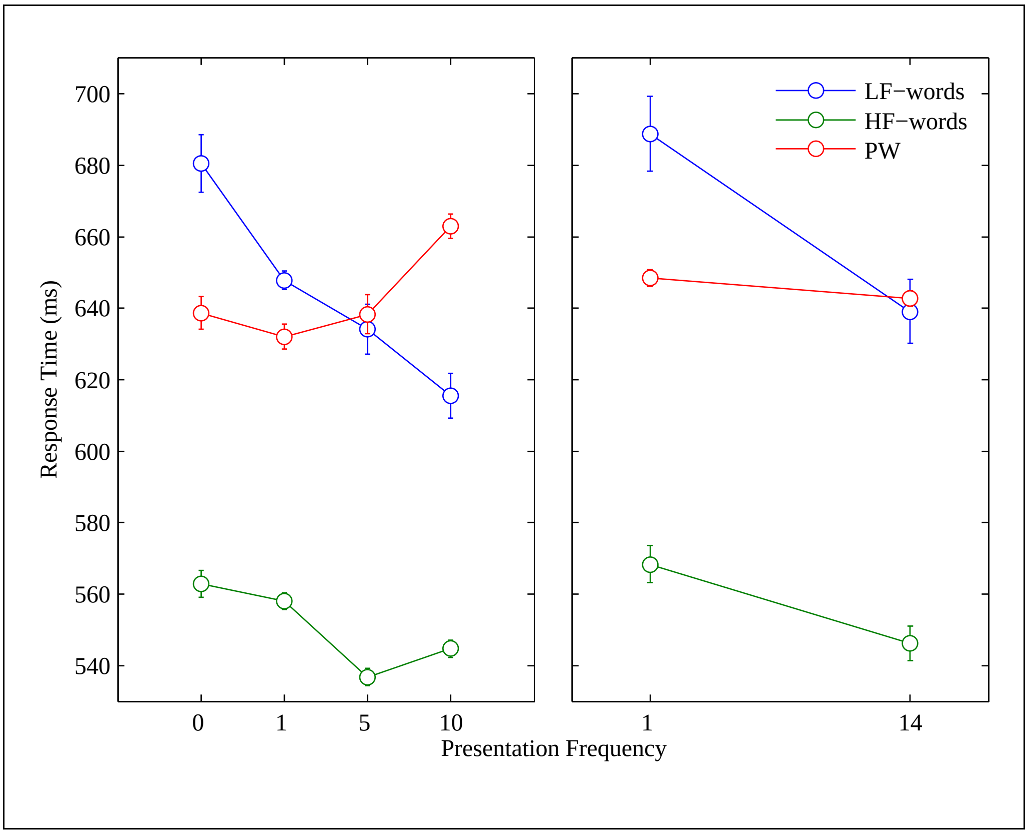

Figure 5: Mean response times of lexical decisions in Experiment 1a for items repeated within word streams (left panel) and across word streams (right panel). LF-words*:* low frequency words, HF-words*:* high frequency words; PW*:* pseudowords. Error bars represent standard error of mean.

Lexical decisions. 6.3% of all responses were incorrect or slow and discarded from analysis. 2.7% of the remaining data were identified as outliers and excluded from further analysis. The left panel of Figure 5 shows mean response time for words and pseudowords repeated within streams as function of presentation frequency. HF-words were responded to faster than LF-words, with averages of 551 ms and 644 ms, respectively. For both, LF-words and HF-words, response times decreased as presentation frequency increased, with an overall effect of 65 ms for LF-words and 18 ms for HF-words.

The LME model with fixed factors for Word Type (LF-words vs. HF-words), Presentation Frequency (0, 1, 5, 10) and their interaction revealed statistically reliable effects of Word Type [$t = 20.50$, $p < 0.001$], Presentation Frequency [$t = 6.27$, $p < 0.001$] and their interaction [$t = -2.03$, $p < 0.05$]. Both random intercepts of Subject and Item proved to be important model terms. The fit became substantially worse when Subject [$\chi^2(1) = 60.97$, $p < 0.001$] or Item [$\chi^2(3) = 756.10$, $p < 0.001$] were excluded.

To investigate if the delay between word stream and lexical decision trial had influenced priming, Trial Number was entered as a continuous predictor variable into the model. Including Trial Number as main effect improved the fit, as revealed by a comparison of the two models [$\chi^2(1) = 4.29$, $p < 0.05$], but did not alter the results for the other fixed effects. Introducing any interaction term including Trial Number did not improve the fit further [all *p*s > 0.2].

Mean response time to pseudowords increased with presentation frequency, from 638 ms for non-presented items to 663 ms for ten times presented items. This effect was reliable [$t = -3.04$, $p < 0.01$]. Dropping one of the intercepts considerably reduced the fit [$\chi^2(1) = 71.5$, $p < 0.001$] and [$\chi^2(1) = 1139.2$, $p < 0.001$]. Including Trial Number as fixed main effect led to the best fitting model [$\chi^2(1) = 12.39$, $p < 0.001$].

Figure 5 (right panel) shows results for items repeated across streams. For LF-words and HF-words mean response time decreased by 50 ms and 22 ms, respectively. In the LME model, the effect of Presentation Frequency was reliable [$t = 3.33$, $p < 0.01$], as well as that of Word Type [$t = 11.17$, $p < 0.01$], but the interaction did not reach significance [$t = -0.78$, $p > 0.40$]. Including Trial Number did not improve the model [$\chi^2(1) = 4.16$, p=0.13]. With regard to random effects, the intercept for Item did not improve the model fit, and was discarded [$\chi^2(3) = 3.6$, $p > 0.30$]. In fact, the estimated variance for Item was essentially zero. Pseudowords did not show any de- or increase in response time [$t = 0.40$, $p > 0.60$], and Trial Number did not contribute to the model fit [$\chi^2(1) = 0.46$, $p = 0.49$]. Nevertheless, discarding intercepts either for Subject or Item worsens the model fit [$\chi^2(1) = 226.8$, $p < 0.001$] and [$\chi^2(1) = 16.1$, $p < 0.001$], respectively.

Impact of stream position. Due to the experimental design and construction of word streams, the position of an item within a stream was highly confounded with its presentation frequency. Although controlled for, mean stream position of items presented 10 times was slightly later than for items presented five times or items presented once. This effect was even exaggerated with regard to the last repetition. Because non-presented items did not have any word stream position, the influence of the position of an item could not be assessed by introducing another predictor into the LME model. Therefore, correlations between word stream position and response times were computed for single participants and conditions. The individual results were then combined using the inverse chi-square method (see Hedges & Olkin, 1985)[1]. Individual correlations ranged from r = -0.32 to r = +0.32 (Figure 6), with significant values for only three correlations. From Figure 6, it is evident that the correlations vary symmetrically around zero (upper panel), and the p-values are distributed uniformly between 0 and 1 (lower panel). Participant-wise analyses across conditions yielded a marginally significant result for participant JG only, [$\chi^2(18) = 28.95$, $p = 0.05$, all other $\chi^2(18) < 20.5$, $p > 0.28$]. Combining p-values across participants yielded marginally significant results for HF-words presented once [$\chi^2(12) = 20.4$, $p = 0.06$] and pseudowords presented ten times [$\chi^2(12) = 18.7$, p=0.10; all other $\chi^2(12) < 17.2$, $p > 0.14$].

[1] Under the null hypothesis every p_i follows a uniform distribution. Given a uniform distribution U, $-2 * \log U$ has a chi-square distribution with two degrees of freedom. Assuming independent p-values, the sum $P = -2 * \Sigma(\log(pi))$ has a chi-square distribution with $2*k$ degrees of freedom and k p-values included. Accordingly, the null hypothesis is rejected if P exceeds the corresponding critical chi-square value.

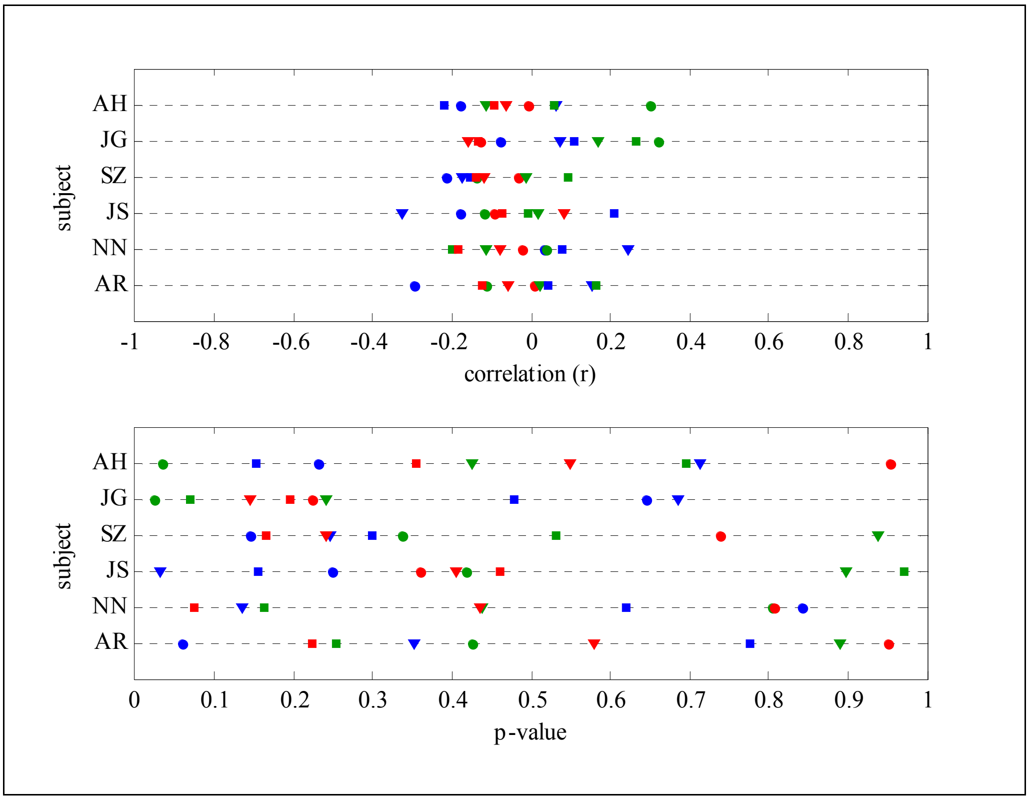

Figure 6: Correlation (upper panel) and corresponding p-values (lower panel) of response times and position of the last word presentation in the preceding word stream for each participant. Colours denote target type: blue – LF-words, green – HF-words, red – pseudowords; symbols denote Presentation Frequency: circle – 1, triangle – 5, square – 10).

Visibility. Visibility was not assessed directly, but performance in the counting task can be used as an indirect measure of how well single items could be consciously perceived in the word stream. Performance in the counting task, assessed by the deviation from the actual number of names that occurred in a stream, was correlated with net priming effects. This gives a hint of whether visibility contributes to priming. This correlation was assessed separately for each word type and presentation frequency on two levels: *Within participants*, counting performance and priming effects of individual runs were correlated. *Across participants*, mean counting performance and mean priming effects for each participant were correlated. Within-participant correlations varied from -0.34 to +.23 (mean -0.04) for LF-words, from -0.35 to +0.27 for HF-words (mean -0.02), and from -0.28 to +0.25 (mean -0.05) for pseudowords, but none reached significance (only 2 out of 54 *ps* < .10, all other *ps* > 0.18). Figure 7 (upper panel) shows that the correlations are symmetrically distributed around zero. The corresponding p-values seem to follow a uniform distribution (Figure

7, lower panel) as confirmed by combination of inverse chi-square values across participants [$\chi^2(12) < 12.2$, $p > 0.38$] and across conditions [$\chi^2(18) < 21.1$, $p > 0.27$].

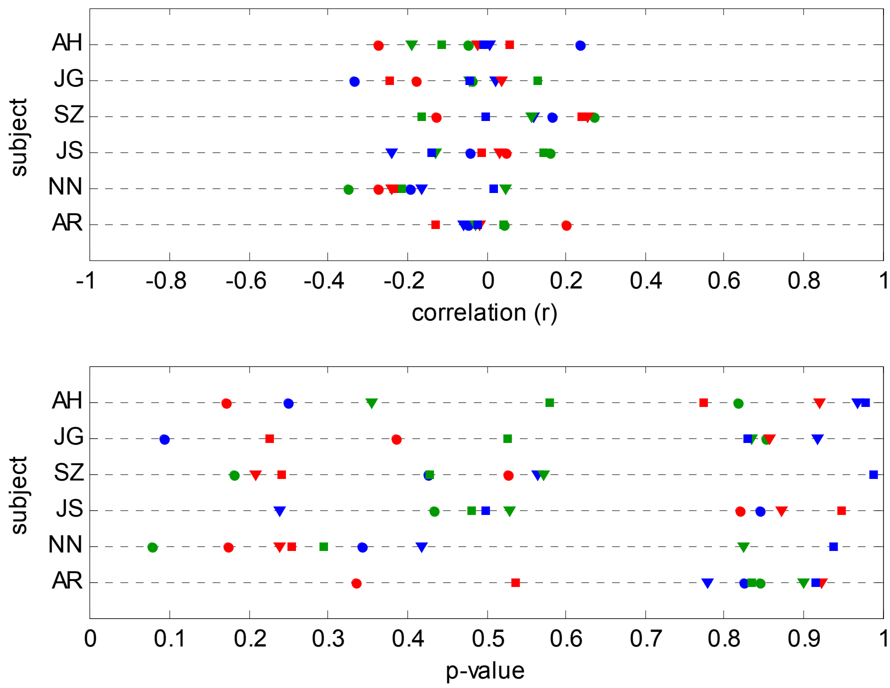

Figure 7: Correlation (upper panel) and corresponding p-values (lower panel) of counting performance and priming effect for each participant. Colours denote Target Type: blue – LF-words, green – HF-words, red – pseudowords; symbols denote Presentation Frequency: circle – 1, triangle – 5, square – 10).

Correlations across participants were similarly inconsistent (Table 4) with just the correlation for pseudowords presented once reaching marginal significance ($r = 0.73$, $p = 0.1$). Note that a positive correlation implies that as performance in the counting task decreases (i.e., the deviation of counted forenames from the actual number increases), facilitating priming increases.

Table 4: Correlations (and p-values) of mean priming effects and mean counting performance across participants. Note, only within repetitions are considered.

	Presentation Frequency		
Type	**1**	**5**	**10**
LF-words	-0,37 *(0,47)*	0,47 *(0,35)*	-0,42 *(0,41)*
HF-words	0,20 *(0,71)*	-0,31 *(0,55)*	0,08 *(0,88)*
Pseudowords	0,73 *(0,10)*	0,22 *(0,67)*	0,12 *(0,82)*

2.1.3 Discussion

The accuracy in the counting task showed that participants focussed their attention on the stream, and points to a good visibility of single items. The lexical decision task revealed substantial priming effects for repetitions within word streams as well as across word streams. Lexical decisions to words were facilitated by repeated presentations within and across word streams. Surprisingly, the effect for repetitions across word streams seemed to be more pronounced than for repetitions within a word stream, which cannot be explained easily by the account outlined above. However, the results for repetitions across word streams have to be met with caution because they rest on a few items only, and thus, some items which generally show large priming effects may bias these results. The same items repeated within a word stream would have not biased the results in the same amount because of more items per condition. This effect normally is controlled for statistically by introducing a random factor Item into the LME model, but in this case the estimation algorithm provided a variance very close to zero, which can be explained by the small number of items in each condition. This may also explain why the interaction of Frequency Type and Presentation Frequency did not reach significance.

Lexical decisions on pseudowords were inhibited when repeated within a word stream, but did not differ when repeated between word streams. Thus, the repetition effects obtained for words and pseudowords showed a different time course, which is reasonable under the assumption that specific lexical activity induced by pseudowords is only weak because no specific representation exists.

Importantly to note is that repetition effects do not depend on the word stream position or the time lag between presentation and lexical decision. The only impact of time lag is an overall slowing of responses with increasing trial number.

The purpose of Experiment 1a was to demonstrate cumulative priming effects by varying the presentation frequency of rapidly presented items for words and pseudowords. The results clearly show that the “word shower” is an appropriate method to produce repetition priming simultaneously in many items. The effects for words are in line with previous findings on repetition priming with a greater amount of priming in low frequency than in high frequency words. The inhibiting effects for pseudowords are in line with findings from Wagenmakers et al. (2004), who argued that inhibitory repetition priming of pseudowords is based on a sense of familiarity that is due to a recent encounter of the pseudoword. Zeelenberg et al. (2004) have shown that the direction of repetition priming of pseudowords depends on the type of study task at hand. Only when the study task as well as the test task is lexical decision, facilitatory effects occur. The study task in the present study was name counting. Because the recognition of a name implies a kind of lexical decision, participants may have performed lexical decisions on each single word stream item. However, the inhibitory effects imply that participants did not engage in lexical decision during the word stream. In sum, according to the approach outlined above, presentation of pseudowords in a word stream leads to activation accumulation in neighbour word representations, and thus, leads to a sense of familiarity and prolonged ‘nonword’ decision times.

Notably, there is no systematic covariation of performance in the name counting task and priming effects, implying that the ability to consciously perceive and process singular items within a word stream does not lead to improved priming. However, this interpretation is somewhat speculative because the counting task is only an indirect measure of visibility and, with an effective processing time of 250 ms for each item of a word stream, there is plenty of time to perceive every single item. Therefore, the next experiment is designed to test, whether the priming effects still hold, when processing time is further decreased.

2.2 Experiment 1b

In Experiment 1a reliable repetition effects were observed within and between word streams. With a presentation duration for each item of 150 ms followed by a blank of 100 ms, there was sufficient time to process each individual item. Thus, in Experiment 1b it was tested whether these stable effects would hold even in case of severely degraded item processing. It is a replication of Experiment 1a but with reduced presentation duration of word streams.

2.2.1 Methods

Participants. Eight students (one male; age 18 - 26 years, mean 20.8 years) from Technical University Braunschweig were tested in two one-hour sessions with at least 2 days apart. All were native German speakers and took part for course credit. Vision was normal or corrected to normal.

Task, Design. See Experiment 1a.

Procedure. The same as in Experiment 1a, except that each word stream item was presented for 54 ms followed by a blank screen for 26 ms, so that effective processing time was 80 ms. Overall word stream duration was 17.1 seconds.

2.2.2 Results

Counting task. The mean counted number of forenames ranges between 4 and 7 items with a mean of 5 names, i.e., participants missed up to 6 forenames in a word stream.

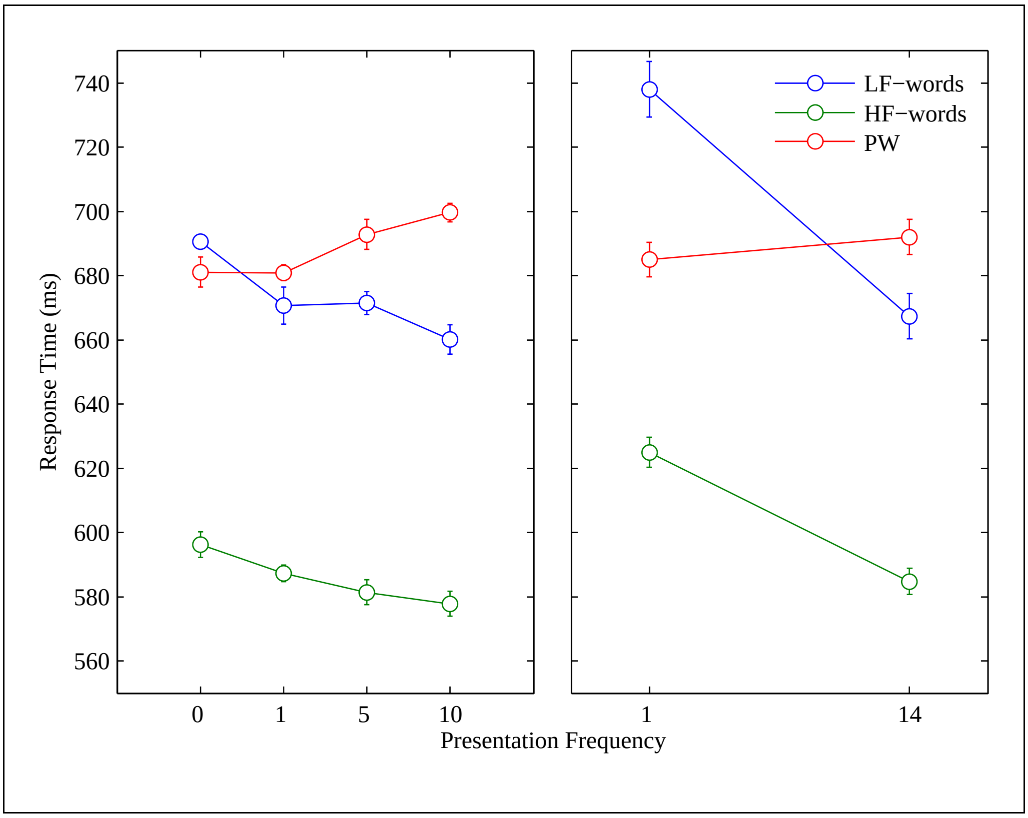

Figure 8: Mean response times of lexical decisions in Experiment 1b for items repeated within word streams (left panel) and across word streams (right panel). LF-words*:* low frequency words, HF-words*:* high frequency words; PW*:* pseudowords. Error bars represent standard error of mean.

Lexical decision task. 7.8% of all responses were incorrect or slow and were discarded. From the remaining trials 2.6% were identified as outliers and were also excluded. Figure 8 (left panel) shows mean response time for words and pseudowords repeated within word streams. For both LF-words and HF-words response time decreased with increasing presentation frequency (30 ms for LF-words and 18 ms for HF-words), with HF-words responded to faster than LF-words (585 ms vs. 673 ms). In the LME model the effect of Word Type and the effect of Presentation Frequency proved to be significant [t = 9.47, p < 0.001 and t = 3.40, p = 0.001, respectively]. Although the effect of Presentation Frequency for LF-words was greater than for HF-words, the interaction did not reach significance [t = -0.50, p > 0.6]. Including Trial Number as main effect did not improve the model fit significantly [$\chi^2(1) = 2.33$, p = 0.13], and no interaction term including Trial Number contributed to the fit [all p > 0.28]. Both random intercepts for Subject and Item were substantial for the model fit.

Dropping either one worsened the model fit [$\chi^2(2) = 31.0$, $p < 0.001$, and $\chi^2(1) = 61.1$, $p < 0.001$].

Response time to pseudowords increased with increasing presentation frequency, from 680 ms for non-presented items to 699 ms for items presented ten times. A LME model with Presentation Frequency as fixed effect, and Subject and Item as random intercepts, renders this effect statistical reliable [$t = -3.13$, $p < 0.01$]. Inclusion of Trial Number as main effect improved the model fit [$\chi^2(1) = 15.26$, $p < 0.001$], but no further improvement was achieved by including the interaction of Trial Number X Presentation Frequency [$\chi^2(3) = 0.77$, $p = 0.85$].

Figure 8 (right panel) shows the results for repetitions across word streams. Response times to LF-words and HF-words decreased by 71 ms and 40 ms, respectively. Introduced in a LME model, the effect of Presentation Frequency was reliable [$t = 4.17$, $p < 0.001$] as well as the effect of Word Type [$t = 5.95$, $p = 0.004$), but no interaction could be observed [$t = -0.78$, $p > 0.40$]. Dropping any random term resulted in a worse fit [$\chi^2(2) = 9.95$, $p < 0.01$ for discarding Subject, and $\chi^2(1) = 13.4$, $p < 0.001$ for discarding Item]. Pseudowords did not show any reliable de- or increase in response time [$t = -0.63$, $p > 0.60$].

Impact of stream position. Within-participant correlations were symmetrically distributed around zero, ranging from $r = -0.53$ [$p = 0.02$] to $r = +0.43$ [$p = 0.07$], with six correlations reaching significance. However, participant-wise combination of p-values yielded a significant result for participant DW only, [$\chi^2(18) = 30.4$, $p = 0.03$], indicating that there is a significant correlation in at least one condition for DW. Combination across conditions did not prove any correlation to be reliable [all $\chi^2(16) < 22.4$, all $p > 0.13$] (Figure 9).

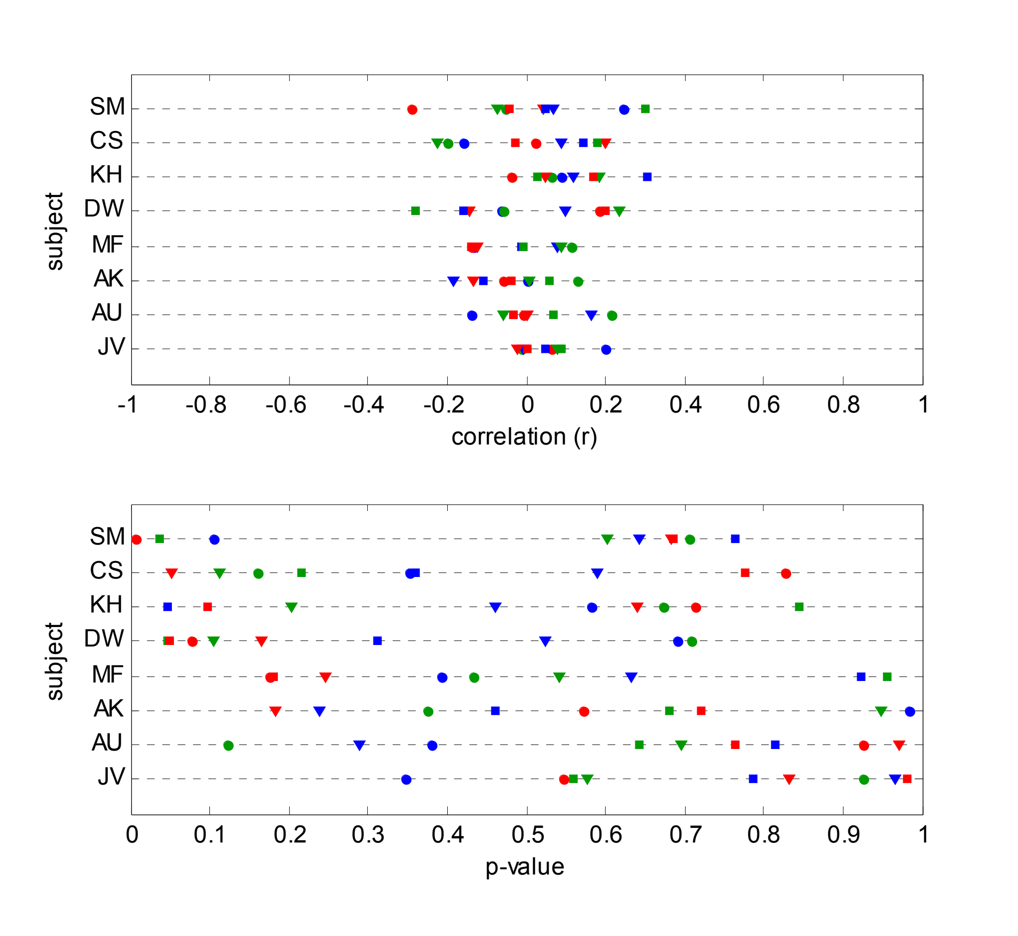

Figure 9: Correlation (upper panel) and corresponding p-values (lower panel) of response times and the position of the last prime presentation in the preceding word stream for each participant. Colours denote Target Type: blue – LF-words, green – HF-words, red – pseudowords; symbols denote Presentation Frequency: circle – 1, triangle – 5, square – 10).

Visibility. Within-participant correlations between counting performance and priming effects varied symmetrically around zero (Figure 10, upper panel), 5 out of 72 correlations reached significance [$p <= 0.05$], (Figure 10, lower panel). Statistical analyses of p-values across conditions revealed a significant deviation from a uniform distribution for participant JV [$\chi^2(18) = 30.4$, $p = 0.03$, all other $\chi^2(18) < 25.2$, $p > 0.12$], indicating that there is at least one condition in which priming effect and counting performance covary. The most reliable correlations for JV are $r = -0.39$ [$p = 0.05$] and $r = -0.60$ [$p < 0.01$] for high frequency words that were presented five times and 10 times, respectively. Note that in the present context a negative correlation indicates a concurrent increase in facilitating priming effects and counting performance.

Combining p-values across participants, HF words presented 10 times deviated significantly from the uniform distribution [$\chi^2(16) = 40.3$, $p < 0.001$, all other $\chi^2(16) < 20.5$, $p > 0.20$]. However, the most reliable correlations for ten times presented HF words did not yield unambiguous evidence, because they differed in direction [participant JV: $r = -0.60$, $p < 0.01$; CS: $r = 0.39$, $p = 0.05$, and SM: $r = 0.44$, $p = 0.02$].

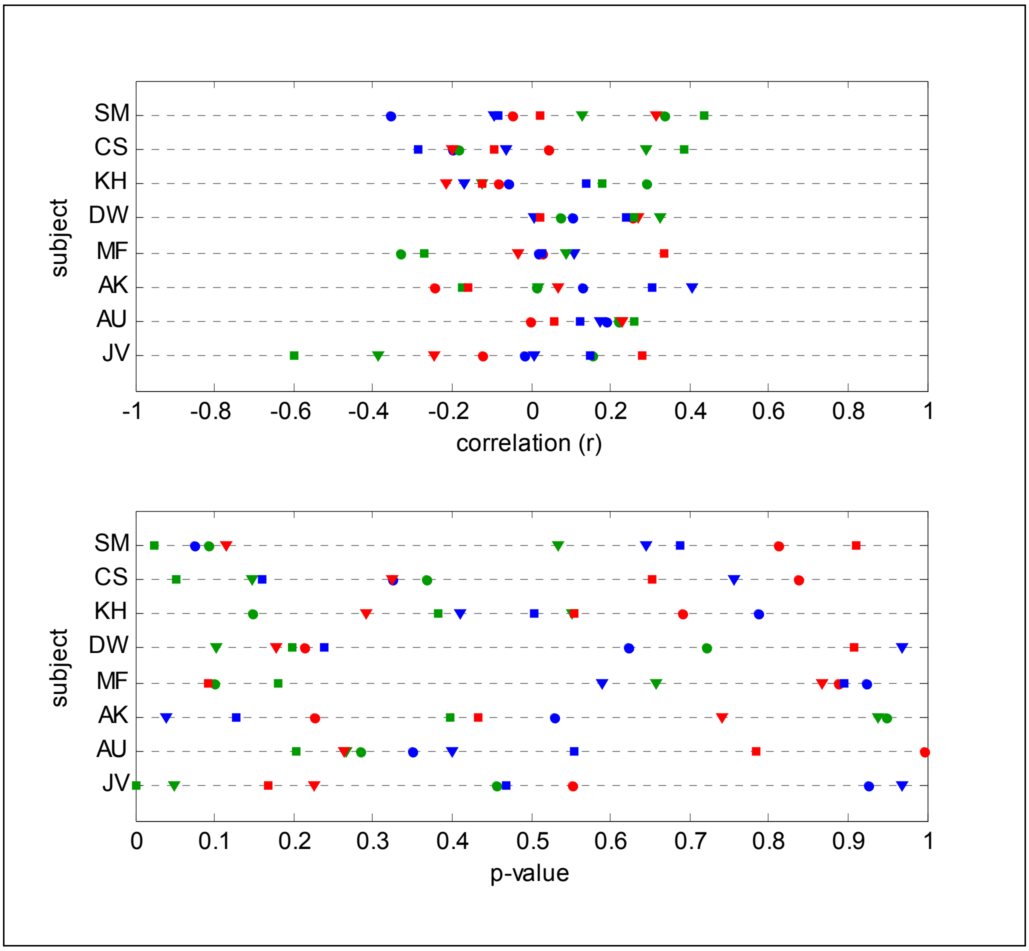

Figure 10: Correlation (upper panel) and corresponding p-values (lower panel) of counting performance and priming effect for each participant. Colours denote Target Type: blue – LF-words, green – HF-words, red – pseudowords; symbols denote Presentation Frequency: circle – 1, triangle – 5, square – 10).

Correlations of mean counting performance and mean priming effects were similarly inconsistent (Table 5), only the correlation for LF words presented five times was marginally significant [$r = -0.64$, $p = 0.09$, all other $p > 0.20$].

Table 5: Correlations (and p-values) of mean priming effects and mean counting performance across participants in Experiment 1b. Note, only within stream repetitions are considered.

	Presentation Frequency		
Type	**1**	**5**	**10**
LF-words	-0,28	-0,64	0,23
	(0,50)	(0,09)	(0,59)
HF-words	-0,32	0,46	-0,45
	(0,44)	(0,25)	(0,26)
Pseudowords	-0,46	-0,48	-0,33
	(0,26)	(0,23)	(0,42)

2.2.3 Discussion

Experiment 1b confirmed and extended the results from Experiment 1a. Performance in the counting task was reduced compared to Experiment 1a, probably reflecting the greater difficulty of the task. It seems that processing of individual items was substantially degraded in most cases.

In lexical decision, reliable facilitation for words was obtained with repetitions within and between word streams. The latter seems at least as strong as for repetitions within streams, but the same objections against an interpretation of differences between repetitions within streams and repetitions across streams as mentioned in Experiment 1a have to be regarded.

Repetitions of pseudowords yielded substantial inhibitory effects within, but no effect between word streams. Thus, it can be concluded that the familiarity of pseudowords increases with more presentations, but that no nonword evidence is encoded during the word stream presentation. The dissociation observed in the comparison of repetitions within and across word stream for words and pseudowords can be explained by MROM: Pseudowords activate word representations only to a small degree, and therefore pseudowords have to be repeated with less items and time intervening than words. Thus, pseudoword repetition yields an effect when repeated within a word stream, but not when repeated across different streams. However, as mentioned above results of repetitions across word streams had to be met cautiously, because the experimental setup did not allow investigating repetitions within and across word steams simultaneously. The effects of repetitions across word streams are considered again in Experiment 6a, but until then the focus is on repetitions within word streams.

Again, additional analyses ruled out the possibility that repetition effects were influenced by word stream position, visibility, and time lag between prime and lexical decision.

Which mechanism underlies the effects for words and pseudowords? Episodic accounts would argue that on each single presentation of a word, the memory trace is enriched and therefore recalled easier and faster at a late test. In the case of repeated nonwords these enriched memory traces could be the basis for a sense of familiarity, which leads to a slowed 'nonword' response. Interactive or abstractionist accounts can explain the effects by a cumulative increase in activation due to repeated presentation, leading to facilitated responses for word targets. Presentation of pseudowords activates the representation of their word neighbours. Thus, the same pseudoword target is responded to slower. A third possible mechanism is based exclusively on experimental factors and needs no assumptions about word or pseudoword processing or any abstract representation of words in a mental lexicon: The probability of perceiving and processing an item at least once increases with presentation frequency. Therefore, the obtained repetition effects might be present only in a subsample of each condition, namely in those items that were fully perceived and processed. Conditions with more repetitions inherit more effect-items than do conditions with only one presentation, and thus, leading to greater mean effects. This issue is dealt with in the next experiment.

3. Experiment 2: Manipulating conscious recollection

Experiments 1a and 1b both showed a reliable repetition effect of words and pseudowords. However, rather than on cumulative activation of word representations, this effect might be based on a higher probability to consciously perceive an item at least once when presented repeatedly.

To assess this question, Experiment 2 manipulates the probability of conscious perception by varying the presentation duration of each item, i.e., the less often a prime was presented in a word stream the longer was its presentation duration on each occurrence. To quantify the probability of perception in each condition a recognition task was introduced in a second session. The rationale beyond this experimental setup is straightforward: Items presented once in a stream are presented for a longer duration than items presented five times or ten times, and should be recognized more accurately in a later recognition task. If the repetition effects are based on a higher probability of conscious perception of individual items, then an item presented once should produce more priming than an item presented more often. However, if priming is based on pure repetition effects, priming should increase with increasing presentation frequency regardless of the presentation duration.

3.1 Methods

Participants. Six students (two male; age 19 - 38 years, mean 24.7 years) from Technical University Braunschweig were tested in two one-hour sessions with at least 2 days apart. All were native German speakers and took part for course credit. Vision was normal or corrected to normal.

Task. In Session 1, participants performed the name counting task and lexical decisions. In Session 2, participants performed the name counting task and a recognition task instead of lexical decisions. They rated individual items whether they had seen them in the preceding stream or not.

Design. Word streams included 266 items with repetitions only within word streams. Due to the longer word stream length, the spacing slightly increased to 13 items. Session 1 (lexical decisions) consisted of 15 runs with one word stream and 64 lexical decisions, each. Session 2 (recognition) consisted of 13 runs with one word stream and 72 recognition trials, each. Presentation Frequency (PF) is treated as categorical factor with levels 0, 1, 5, and 10).

Procedure. Word stream presentation was changed in the following respects: Immediately before the first and after the last word stream item three filler items were presented for 54 ms followed by a blank screen for 26 ms. Filler items consisted of 8 special characters ("########", "%%%%%%%%", "&&&&&&&&"). Presentation duration of each single item was chosen so that the overall presentation durations of words and pseudowords in the different repetition conditions were approximately constant, i.e., items with Presentation Frequency 1, 5, and 10 were presented for 530 ms, 105 ms, and 53 ms, respectively. The recognition task in Session 2 was similar to the lexical decision task in Session 1 except that no time limit and no feedback were given.

Data analysis. Because in this experiment single repetition conditions were of interest instead of an overall data pattern, no linear contrasts are reported, but difference contrasts of adjacent conditions. Recognition data were analyzed by linear mixed effects logistic regression by providing a logit link function. In all other regards the analysis was the same as for response time data.

3.2 Results

Counting task. Mean deviation from the actual number of 10 names was m = -2.46 (SD = 0.66, median = 2), i.e., on average participants missed slightly more than 2 items.

Lexical decision task. 6.2% of all trials were discarded due to incorrect or slow responses. 2.7% of the remaining trials were identified as outliers and discarded, too. Figure 11 (left panel) shows mean response times. For words, response times decreased from PF 0 to PF 1 [t = 3.68, p < 0.001], and from PF 1 to PF 5 [t = 2.73, p=0.006], but did not further decrease from PF 5 to PF 10 [t = -1.34, p = 0.18]. The effect for Word Type was also highly significant [t = 11.99, p = 0.001], as well as the interaction of Word Type with PF 0 vs. PF 1 [t = -2.34, p = 0.02]. All other interactions were not reliable [|t| < 1.6, p > 0.10].

Response Time for pseudowords showed the complementary pattern to words, in that they increased from PF 0 to PF 5, and then decrease to PF 10. But only the difference between PF 1 and PF 5 proved to be reliable [t = -2.16, p = 0.03; both other |t| < 1.3, p > 0.2].

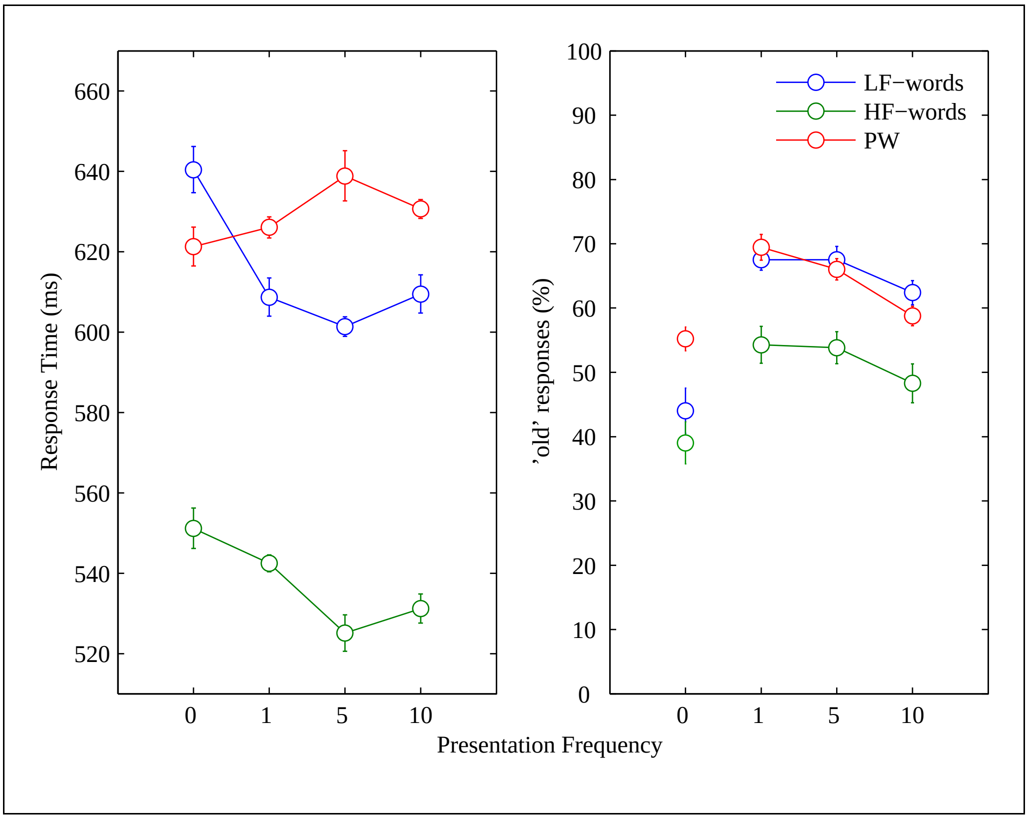

Figure 11: Mean response time of lexical decisions (left panel) and recognition performance (right panel) for Experiment 2. Error bars represent standard error of mean.

Recognition task. The right panel of Figure 11 shows mean percentage of 'old' responses. LF-words were better recognized than HF-words, 63.2 % vs. 54.1% [z = 4.38, p < 0.001]. Of most interest was the effect of Presentation Frequency on recognition rate: For words, percentage correct were 58.3%, 60.5%, 60.8%, and 55.1% for PF 0, 1, 5, and 10, respectively. Only the difference between PF 5 and PF 10 was marginally significant [z = 1.72, p = 0.09; both other $z < 1$, $p > 0.3$]. In addition, the interaction of Word Type and the difference of PF 0 and PF 1 proved to be significant [z = 3.51, $p < 0.001$].

For pseudowords, mean correct rejection rate (PF 0) was 44.5 %. The recognition rate steadily decreased from PF 1 (69.4%) over PF 5 (65.8%) to PF 10 (58.3%). The difference between PF 1 and PF 5 did not reach significance [z = 1.1, $p > 0.2$], but the differences between PF 5 and PF 10 [z = 2.27, $p < 0.05$] and between PF 0 and PF 1 were reliable [z = -9.1, $p < 0.001$].

3.3 Discussion

The results are highly similar to those obtained in Experiments 1a and 1b. Although presentation durations were drastically reduced with increasing presentation frequency, reliable priming modulation by presentation frequency was found in lexical decisions. Words presented five times were faster responded to than words presented once, which in turn were faster responded to than items that had not been presented at all. Response time to words presented ten times did not yield any incremental priming effect.

In contrast, recognition performance did not differ reliably between new words, words presented once, and words presented five times. This clear dissociation leads to the conclusion that the probability of consciously perceiving a word at least once within a word stream (and remembering it later) cannot account for the obtained effects. Nevertheless, there may be a contribution of conscious perception, at least when presentation durations are short compared to the other items in the word stream. This can be seen by the simultaneous decrease in recognition rates and priming effects for items presented ten times.

4. Experiment 3: Quantifying recognition of repeated primes

The aim of Experiment 3 was to further confirm massive repetition priming and to test the hypothesis that the effects depend on the counting task. According to a component-process approach (Vriezen, Moscovitch & Bellos, 1995) long-term priming depends on the highest level of processing required at study and at test. For example, a lexical decision task at study should induce repetition priming when the test task is lexical decision, but not when the test task involves a semantic judgement (e.g. “Decide as fast as possible whether the presented word describes an animal or an object.”), because the highest level of processing required in a semantic judgement is higher than in a lexical decision. Vice versa, a semantic judgement at study should produce repetition priming regardless of whether the test task is a semantic judgement or a lexical decision. Further, several authors demonstrated that repetition priming effects, especially for nonwords, depend on the encoding task (e.g. Zeelenberg et al., 2004; Masson & MacLeod, 2000; Masson & Hicks, 1999). Counting names in a word stream requires a decision about the “word-likeness” and the semantic content of a presented item. Therefore, it might be that repetition effects diminish, when the counting task is more superficial. In the following experiment, the counting task was chosen so that it could be accomplished without elaboration of lexicality or semantics.

A second goal was to quantify the relation of presentation frequency and recognition performance and the covariation with priming effects more thoroughly. As in Experiment 2, the second session consisted of recognition trials instead of lexical decisions.

Further, it has to be noted that the item pool was completely overhauled due to some minor problems. First, interference within and between words and pseudowords was not controlled for, e.g. the words “MUND” and “MOND” both were used as targets which might have introduced some bias or at least error variance. Also, the word “BESEN” and the pseudoword “GESEN” both were used in the preceding experiments. Due to the randomization procedure these neighbours might have occurred in the same experimental block, leading to biased lexical decision latencies. Further, minor flaws in the legality of pseudowords were fixed.

4.1 Methods

Participants. Twelve students (one male; age 20 - 28 years, mean 21.5 years) from the Technical University Braunschweig were tested in two one-hour sessions which were at least 5 days apart. All were native German speakers and took part for course credit. Vision was normal or corrected to normal.

Task. Participants were instructed to monitor the number of words written in capitals. No information was given that all these words were forenames. After each word stream a block of lexical decisions (Session 1) or a block of recognition trials (Session 2) followed. At the end of Session 2 participants were asked whether they recognized that all counting items had been forenames.

Stimuli. 512 LF-words and 512 HF-words with frequency counts of less than 5 per million (mean 1.4) and more than 9 per million (mean 69.0), respectively, were used as stimuli. All words were 3 to 6 letters long. Direct neighbourhood was limited to a minimum: only 26 words had one direct neighbour within the set of words. Pseudowords were constructed with one "parent" word in the word stimulus set, so that each pseudoword had exactly one direct neighbour in the word set. Thus, pseudowords could be classified as "low frequency" and "high frequency" pseudowords depending on the frequency of their "parent". Allocation of stimuli to experimental conditions was counterbalanced across participants and conditions, so that each participant never got a word and his neighbour pseudoword in the same session. For details on stimulus characteristics see Table 6 and 7. Counting items were the same names as in all other experiments typed in capitals.

Table 6: Stimulus characteristics in Experiments 3-5

	Letters mean (SD)	Syllables mean (SD)	CELEX frequency count/million	CELEX frequency log/million	Trigram frequency
LF-words	5.3 (0.8)	1.7 (0.5)	1.4 (0.8)	0.17 (0.20)	157.8 (20.0)
HF-words	5.2 (0.8)	1.6 (0.5)	69.0 (139.5)	1.56 (0.44)	162.5 (19.8)
LF-pseudowords	5.2 (0.8)	1.7 (0.5)	--	--	160.6 (19.4)
HF-pseudowords	5.1 (0.8)	1.6 (0.5)			162.9 (19.7)

Table 7: Mean letter overlap within and between stimuli classes in Experiments 3-5

	LF-words	HF-words	LF-pseudowords	HF-pseudowords
LF-words	0.29 (0.19)	0.29 (0.19)	0.30 (0.19)	0.29 (0.19)
HF-words		0.28 (0.19)	0.29 (0.19)	0.29 (0.19)
LF-pseudowords			0.32 (0.20)	0.32 (0.20)
HF-pseudowords				0.31 (0.20)

Design. Session 1 consisted of 16 runs with one word stream of 266 items each, followed by 64 lexical decisions. Session 2 consisted of 13 runs with one word stream of 266 items each, followed by 72 recognition trials. Independent variables were item type (LF-words, HF-words, LF-pseudowords, HF-pseudowords), and Presentation Frequency (0, 1, 5, 10). Main dependent variable was Response Time in lexical decision trials (Session 1) and Accuracy in recognition trials (Session 2).

Procedure. The procedure was the same as in Experiment 2 with the exception that all word stream items were presented for 54 ms followed by a blank screen for 26 ms.

4.2 Results

Counting task. Mean deviation from actual number was -2.2 items (SD = 1.9, median = 2.5), i.e., participants missed some capital letter words. None of the participants had recognized that counting items had been forenames.

Lexical decision. 9.4% of all responses were incorrect or slow. 1.7% of the remaining trials were identified as outliers. All these trials were excluded from further analyses. Mean response time of lexical decisions is shown in Figure 12 (left panel). Mean response times for LF-words and HF-words were 656 ms and 580 ms, respectively [t =11.17, p < 0.001]. With increasing presentation frequency response times decreased from 630 ms for new items to 607 ms for items presented ten times. This linear trend was statistically reliable [t = 6.08, p < 0.001]. No interaction of Word Type and Presentation Frequency was observed [t = -1.19, p = 0.23].

Mean response times for LF-pseudowords and HF-pseudowords were 684 ms and 691 ms, respectively. Although minimal in amount, this difference was marginal significant [t = -2.02, p=0.05]. With increasing presentation frequency, response times increased from 684 ms up to 694 ms [t =-2.48, p < 0.02]. No interaction was observed [t = 0.79, p > 0.40].

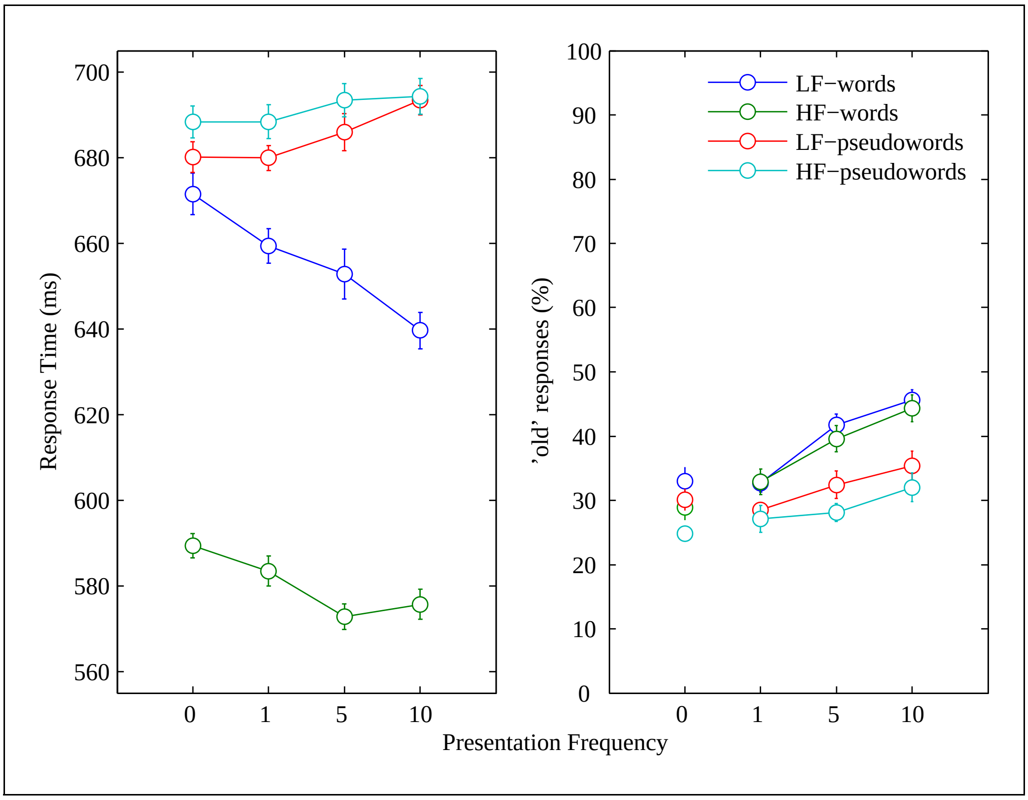

Figure 12: Mean response time of lexical decisions (left panel) and percentage of old-responses (right panel) for Experiment 3. Error bars represent standard error of mean.

Recognition task. Mean percentage of ‘old’ responses to words and pseudowords is shown in Figure 12 (right panel). The frequency of false alarms (i.e., ‘old’ responses to new items) was similar for words (33.0% and 28.9% for LF-words and HF-words, respectively) and pseudowords (30.1% and 24.9.1% for LF-pseudowords and HF-pseudowords, respectively). Entered into a logistic regression LME model the contrast between LF-pseudowords and HF-pseudowords reached significance [z =-3.2, $p < 0.01$], whereas the difference between LF-words and HF-words did not reach significance (z = -1.6, p = 0.11]. Percentage of words and pseudowords correctly recognized as old increased with presentation frequency from 32.6% to 45.6% (LF-words), 32.9% to 44.4% (HF-words), 28.5% to 35.4% (LF-pseudowords), and 27.1% to 32.0% (HF-pseudowords). Data from “old”-trials were entered into a logistic regression LME model separately for words and pseudowords. There was a significant linear trend of Presentation Frequency for words (z =-5.44, $p < 0.001$), but no effect of Word Type nor an interaction of Presentation Frequency and Word Type (both $z < 1$, both $p > 0.6$). Pseudowords showed exactly the same pattern: a reliable linear trend of Presentation Frequency (z = -2.86, p = 0.004) but no reliable effect of Type [z = 1.76, p = 0.08] and no interaction [z = 0.43, p = 0.67]. For purposes of illustration, recognition performance is given in terms of signal detection theory’s *d’* in Table 8. It clearly shows that recognition performance increased with presentation frequency, but was rather poor even for items presented ten times (d’ < 0.5).

Table 8: Mean recognition performance (d’) of words and pseudowords (Session 2).

Target Type	Presentation Frequency		
	1x	5x	10x
LF-words	-0.04	0.28	0.38
HF-words	0.14	0.34	0.46
LF-pseudowords	-0.09	0.07	0.15
HF-pseudowords	0.13	0.08	0.28

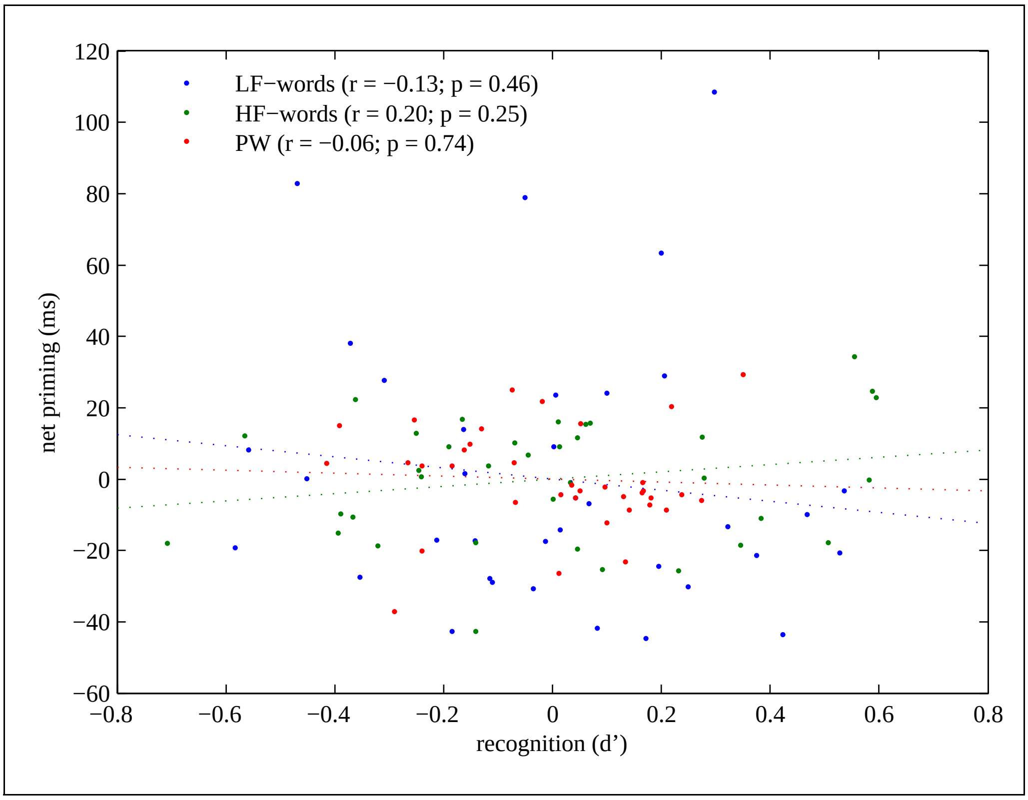

Figure 13: Covariation of recognition performance (d') and priming effects (ms) in Experiment 3.

Recognition vs. Priming. To assess the covariation of repetition priming with conscious recollection, the net priming effects were correlated with *d'*. Individual net priming effects were assessed by the difference of individual mean response times to new items and individual mean response times of Presentation Frequency 1, 5, and 10, respectively. Then, the mean net priming and mean recognition performance was set to zero in all conditions, so that covariation of the amount of priming and the amount of recognition cannot be attributed to different experimental conditions, but can be interpreted as direct relation between recognition and priming. Data on pseudowords were collapsed over Word Type, so that no distinction between LF-pseudowords and HF-pseudowords was made. Figure 13 shows the resulting scatterplot. For all item types correlations were $r < 0.2$, and did not prove statistically reliable [all *ps* > 0.25].

4.3 Discussion

The results show that stable repetition priming occurs with an attentional task that does not require processing of semantic information like the name counting task used so far, but solely relies on analysis of visual features. Performance in the counting task was not perfect but considerably better than in Experiment 1b. Further, the fact that participants did not recognize the counted items as forenames corroborates that participants did not engage in lexical processing. Although the counting task in the present experiment did not require processing of word content or word-likeness, lexical decisions to words were facilitated and lexical decisions to pseudowords were inhibited. Further, responses to LF-pseudowords were slightly faster than responses to HF-pseudowords. However, this priming effect is so small in amount that it has to be confirmed further.

Repetition of words and pseudowords also influenced the recognition of words and pseudowords. For both words and pseudowords recognition performance increased with presentation frequency. Thus, words presented more often are better recognized as old and faster categorized as words, whereas pseudowords are better recognized as old but categorized slower as pseudowords. This is what one expects, when decisions are based on familiarity as Wagenmakers et al. (2004) concluded. If an item is familiar in lexical decision, one would tend to respond “word” and therefore would be faster on repeated words and slower on repeated pseudowords. On the other hand, on recognition trials familiar items trigger the response “old” regardless of item type. But what mechanism underlies these effects? The covariation of recognition and lexical decision seems to imply a direct connection and a common basis for both effects.

Nevertheless, a closer look at the correlations between priming effects and recognition performance reveals no covariation beyond presentation frequency. Presentation frequency influences both recognition and lexical decision, but there seems no causal link between the latter two. This conclusion is in line with Ratcliff et al. (1985), who showed different time courses for repetition priming and recognition memory in long-term priming. Thus, by massive repetition of words and pseudowords, priming is produced that is independent of their conscious recollection at test. In terms of MROM, these results indicate that the activation of an abstract word representation does not necessarily lead to the formation of an episodic memory trace which can be recollected later. Rather, the lexical representation is changed by altering the values of one or several decision thresholds without affecting episodic memory.

5. Experiment 4: The longevity of massive priming

The aim of Experiment 4 was to test the temporal limits of massive repetition priming. In long-term priming, priming effects over several minutes can be observed (e.g. Wagenmakers et al., 2004), whereas in conditions, in which the prime is viewed passively for a short duration, only a short lived effect is found (e.g. Forster & Davies, 1984). In their influential paper Ratcliff et al. (1985) attribute these different priming time-courses to two different components of activation: a relatively short-lived and decaying component, and an almost constant long-term component. In the experiments presented so far, the mean time span between presentation and lexical decision was already in the minute range. If effects rely on transient activity, one would expect them to diminish over time, whereas if effects are based on a constant component, manipulating the delay between presentation and test should have no influence. Therefore, in the next experiment the time lag between presentation and lexical decision was manipulated.

5.1 Methods

Participants. 13 students (one male; age 19 - 41 years, mean 25.0 years) from Technical University Braunschweig were tested in two one-hour sessions with at least 2 days apart. All were native German speakers and took part for course credit. Vision was normal or corrected to normal.

Task, Stimuli. Task and Stimuli were the same as in Experiment 3.

Design. Each session consisted of 14 runs with one word stream of 266 items each, followed by 64 lexical decisions. Items were repeated within word streams only. For half of the participants, half of the word-pseudoword pairs were assigned word targets in Session 1 and pseudoword targets in Session 2, and the other half vice versa. So word and pseudoword neighbours never occurred in the same session. Item allocation to delay and repetition conditions was random.

Independent variables were Item Type (LF-words, HF-words, LF-pseudowords, HF-pseudowords), Presentation Frequency (0, 1, 5, 10) and Delay (immediate test, delayed test). One half of all items in a stream were tested immediately, i.e., in same run. The other half was tested with delay, i.e., in the next run. Note, that “immediately tested words” were tested after approximately 2 minutes, whereas “delayed tested words” were tested after more than 7 minutes. Because in the first run of a session no “next block” items were available for test, “next block” items of the last block were tested in lexical decision without being presented before. So the number of lexical decisions was held constant across all runs. Data from the first block of each session were excluded from all analyses. Lexical decision response time was the main dependent variable.

5.2 Results.

Counting task. On average participants missed 4.1 of the 10 forenames (SD = 2.1, median = 5).

Lexical decision task. 5.3% of all responses were incorrect or slow, and 1.7% of the remaining trials were identified as outliers. These trials were excluded form further analyses. Figure 14 shows the results for the lexical decision task. Overall, the same pattern as in the preceding experiments was observed. HF-words were responded to faster than LF-words [557 ms vs. 628 ms; t = 15.45, p < 0.001], and response time to words decreased with increasing Presentation Frequency [t = 6.85, p < 0.001]. The interaction between Word Type and Presentation Frequency did not reach significance [t = -1.15, p = 0.24]. As can be seen from the right and left panels of Figure 14, data for immediate test and delayed test virtually look the same, neither the main effect nor any interaction involving Delay as factor were reliable [all p > 0.60].

Pseudowords showed a reliable effect of Presentation Frequency with response time increasing from 645 ms for new items up to 660 ms for items presented ten times [t = -5.26, p < 0.001]. LF-pseudowords were responded to slightly slower than were high frequency pseudowords (654 ms vs. 648 ms). This effect was marginally significant [t = 2.1, p = 0.05]. In addition, the interaction of Item Type and Delay was reliable [t = -2.22, p = 0.03], as response time minimally decreased from immediate test to delayed test for LF- pseudowords, but minimally increased for HF-pseudowords. All other effects turned out not reliable [all ps>0.15].

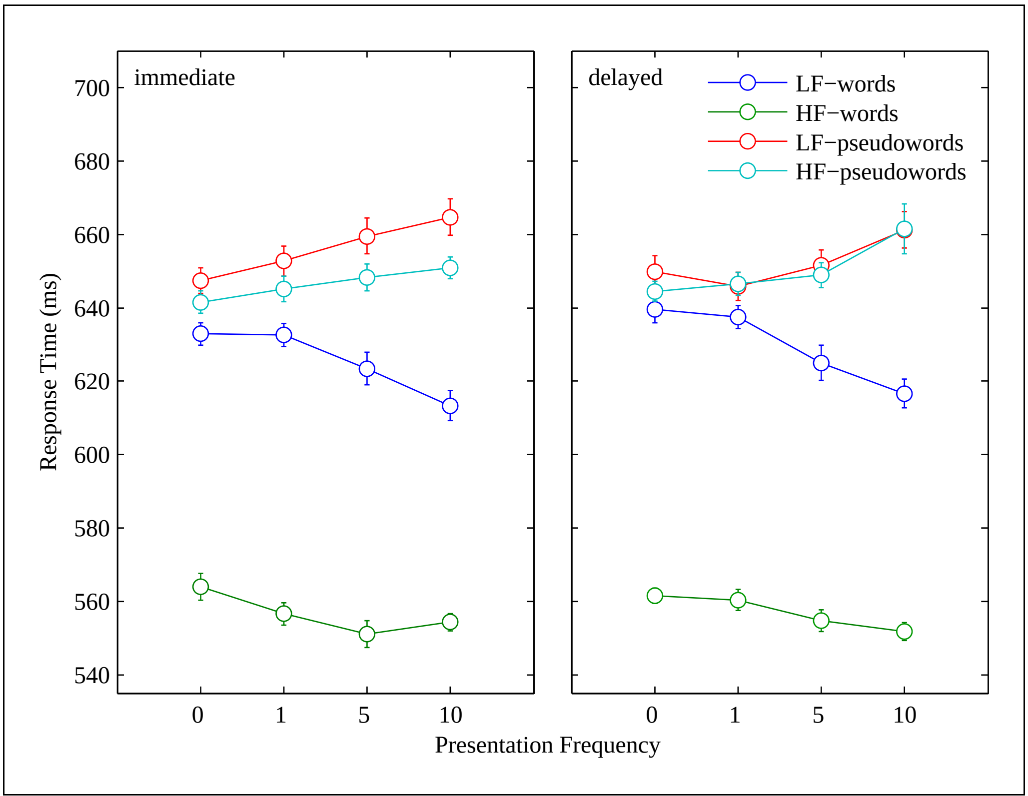

Figure 14: Mean response time of lexical decisions in Experiment 3. Immediately tested items were presented and tested in the same run (left panel), delayed tested items were presented and tested in successive runs (right panel). Error bars represent standard error of mean.

5.3 Discussion

The present experiment investigated the longevity of repetition priming effects of words and pseudowords presented in rapidly presented lists. Lexical decisions on words were facilitated with increasing number of occurrences in the word stream, whereas the processing of pseudowords was inhibited. Remarkably, effects were of the same amount whether tested immediately after the word stream in which targets were presented or after the next word stream. Note that the term “immediately tested words” refers to a mean delay of approximately 2 minutes between presentation in the word stream and lexical decision. The delay between presentation and the lexical decision after the following block add up to more than 7 minutes. These results clearly indicate that the repetition effect measured in the presented experiments is based on a stable long-term component (cf. Ratcliff et al., 1985, McKone, 1995). The different amounts of priming obtained by varying the presentation frequency on immediate test may be explained by different levels of transient activity. However, the fact that the amount of priming stays the same after another couple of minutes indicates that activation does not decay but remains on the same level for several minutes.

Further, results indicate that there may be differences in pseudoword processing based on the frequency class of the parent word. Pseudowords created from LF-words yielded a slightly slower response than did pseudowords created from HF-words. This contradicts the predictions of MROM and the results of Experiment 3, in which responses were faster to LF-pseudowords than responses to HF-pseudowords. However, the effects found in the present study were small and obtained only on immediate test.

6. Experiment 5: Cross-priming of words and pseudowords

Experiments 1 to 4 yielded stable evidence for incremental priming effects of repeated items in rapidly presented word streams. According to MROM and REM-LD, presentation of pseudowords leads to activation/retrieval of the representation of neighbour words due to their similarity. Thus, these models in fact predict cross-priming of words and pseudowords, both of which are based on word representations. To test this hypothesis, for half of the tested words and pseudowords I presented the corresponding neighbour in the word stream. I hypothesized that pseudowords facilitate the following lexical decision on words, but to a lesser amount than do repetition word primes. On the other hand, word primes should enhance inhibitory priming on pseudowords.

6.1 Methods

Participants. Sixteen students (three male; age 19 - 36 years, mean 22.3 years) from Technical University Braunschweig were tested in two one-hour sessions with at least 2 days apart. All were native German speakers and took part for course credit. Vision was normal or corrected to normal.

Task, Stimuli, Procedure. Task, Stimuli and Procedure were the same as in Experiment 3.

Design. Each session consisted of 14 runs with one word stream of 266 items, each followed by 64 lexical decisions. Items were repeated within word streams only. Independent variables were Word Type (LF-words, HF-words, LF-pseudowords, HF-pseudowords), Presentation Frequency (0, 1, 5, 10) and Prime Type (repetition, neighbour). On repetition trials, the same item was tested that had been presented in the word stream, on neighbour trials the corresponding word neighbour had been presented for pseudoword targets, and the pseudoword neighbour for word targets.

6.2 Results

Counting task. Mean number of counted names was 5.6 (SD = 2.4) names per word stream, mean deviation was -4.4, i.e., participants missed approximately half of the names presented.

Lexical decision task. 7.2% of all trials that were incorrect or slow, and further 1.3% of the remaining trials identified as outliers were excluded from analyses. Figure 15 depicts mean lexical decision times to words and pseudowords for each condition. First, the typical pattern for words was observed: HF-words were responded to faster than LF-words [563 ms vs. 629 ms; t = 22.51, p < 0.001], and response time decreased from 602 ms to 588 ms [t = 7.52, p < 0.001] from Presentation Frequency 0 to Presentation Frequency 10. Again, these effects were additive as the interaction did not reach significance [t = 1.01, p > 0.3]. As can be seen from Figure 15, the data pattern looks similar for repetition and neighbour trials. Nevertheless, the effect of Prime Type and the interaction of Prime Type and Presentation Frequency were both reliable, [t =2.97, p < 0.01 and t =2.72, p < 0.01, respectively]. Repetition facilitation was greater on repetition trials (606 ms to 582 ms) than on neighbour trials (600 ms to 593 ms).

Lexical decision latencies for LF-pseudowords and HF-pseudowords did not differ reliably [656 ms vs. 657 ms, t =-0.46, p > 0.6], but increased slightly with increasing Presentation Frequency [654 ms to 660 ms, t =-2.48, p = 0.02]. The interaction of Presentation Frequency and Pseudoword Type did not prove statistically reliable [t = 1.62, p = 0.11]. There was a marginal significant interaction of Prime Type and Pseudoword Type [t =1.82, p = 0.09] and no interaction of Prime Type and Presentation Frequency [t =-1.26, p > 0.2], although the linear trend was slightly greater and more consistent on repetition trials. No other reliable effects were observed [all ps > 0.3].

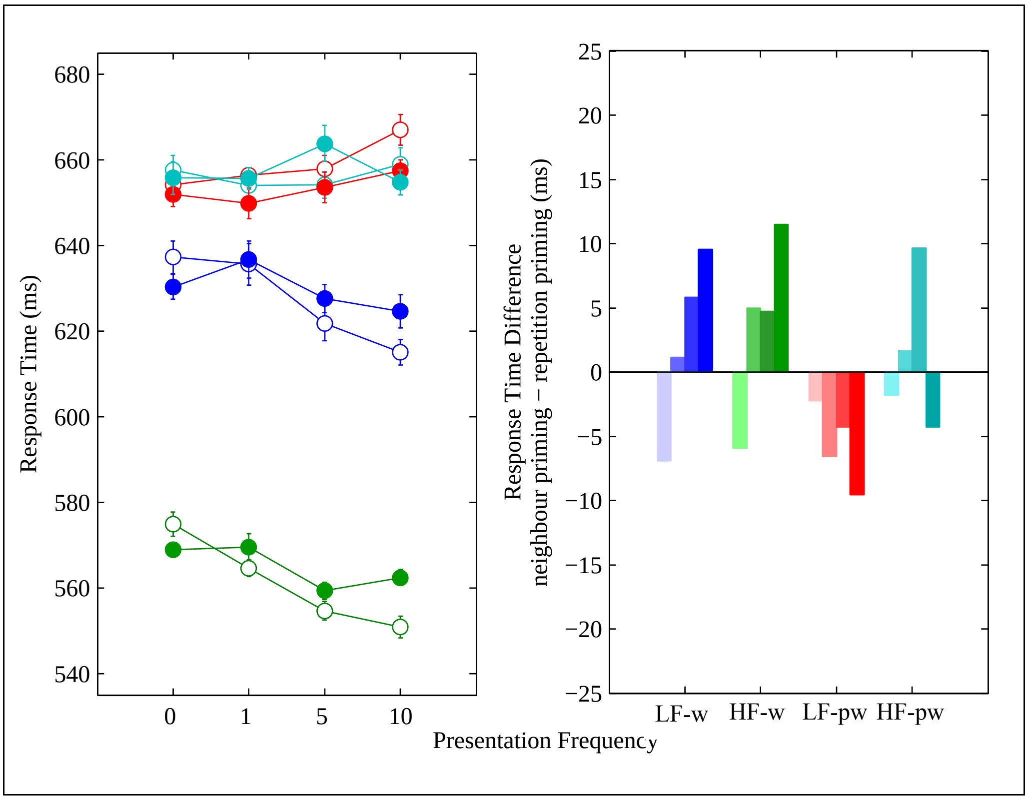

Figure 15: *Left panel:* mean response times in Experiment 5 for repetition trials (open circles) and neighbour trials (closed circles); blue: LF-words, green: HF-words, red: LF-pseudowords, turquoise: HF-pseudowords. *Right panel:* difference effects of neighbour prime – repetition prime. Shading of bars indicate the different Presentation Frequencies from newly presented items (lightest bars) to items presented ten times (darkest bars); w – words, pw – pseudowords. Error bars represent standard error of mean.

6.3 Discussion

Experiment 5 replicated the findings found so far. Increasing presentation frequency increased facilitation of lexical decisions on words and increased inhibition of lexical decisions on pseudowords. Again, repetition effects for words were not modulated by word frequency, and, in contrast to Experiment 4, no difference in the processing of LF-pseudowords and HF-pseudowords was found.

For words, repetition primes induced slightly more facilitation than did pseudoword neighbour primes. This finding is consistent with the assumption of activation of word representations by pseudowords (Wagenmakers et al., 2004; Bölte, 2001; Grainger & Jacobs, 1996). A pseudoword activates the representation of a neighbour word less than that of the word itself, and thus smaller positive priming effects are observed on word trials. For pseudowords, no reliable difference was found between repetition primes and word neighbour primes. According to the present account, enhanced priming should have been observed because lexical activity is higher on presentation of the word prime than on presentation of the pseudoword, and therefore should delay the “nonword”-response more. Therefore it seems plausible that inhibitory pseudoword priming and facilitatory word priming depend on different mechanisms.

In contrast to Experiment 3, Experiment 4 and to the findings of Perea et al. (2005), LF-pseudowords showed the same data pattern as HF-pseudowords. However, a problem of the present study was that pseudowords were controlled for neighbours in the stimulus set but not for neighbours in general. Therefore, it may be that effects are augmented by the total number of neighbours. On the other hand, even Perea et al. (2005) obtained effects only for the fastest 10% of all responses. Thus, the effects of pseudoword type remain ambiguous and do not suggest an interpretation.

7. Experiment 6a and 6b: Short streams and long lags

7.1 Experiment 6a

The experiments so far gave ample evidence of massive repetition priming within rapidly presented word streams. The effects for words and pseudowords were stable over a variety of conditions, including neighbour priming (Experiment 5), in which word neighbour primes resulted in slowed responses to pseudowords, and pseudoword neighbour primes resulted in speeded responses on words, although the effects were considerable smaller than repetition effects. To refine the method depicted so far, in the following experiment word streams were shortened to 5-6 seconds (42-58 items), with one target item per word stream, which was tested immediately after the relevant word stream. With word streams that short, items were either presented once or five times, or were not presented at all. With better control over targets and possible effective primes, more profound and stable repetition effects are expected.

In research on the acquisition of language and words, it is discussed whether repeated exposure to nonwords is sufficient to transform them into words, i.e., to create an entry in the mental lexicon. There is some evidence that nonwords repeatedly encountered over days and weeks yield response patterns typical for processing of words (mirror effects, frequency effects). For example, Gaskell and Dumay (2003) presented nonwords on day one. On day two, these nonwords inhibited the processing of neighbour words. The authors take this as evidence for a increasing word-likeness of nonwords, when participants had processed them. In these studies participants gave overt responses to pseudowords on both days. But the question remains whether the development of lexical characteristics is possible by multiple presentations of pseudowords within word streams, when the presented nonwords are non-targets on day one, and are presented so rapidly that processing of each individual nonword is degraded.

For words, long-term repetition priming lasting several days or weeks (and in some cases even years) is widely acknowledged (e.g. van Turennout et al., 2000; Cave, 1997). In these studies participants always processed items as targets in an encoding task (e.g. lexical decision or naming). However, in light of the present experiments the question remains whether items presented as distractors in a word stream may have an impact that exceeds the momentary experimental context. To recapitulate, in Experiment 4 it was shown that the effects hold at least for several minutes (one word stream + one lexical decision block). In the present experiment, effects over 24 hours are tested for items that had been presented either as distractors in word streams (D-targets) or as lexical decision targets (T-targets). Thus, it is possible to disentangle effects of episodic memory and pure repetition. On T-target trials repetition effects may be due to the specific episode of making a lexical decision on this item. On D-Target trials it is less probable that there is an episodic memory of this particular item. So, effects in this condition would indicate pure repetition effects. For pseudowords, different predictions have to be made. For T-targets, there should be no inhibiting effect because these targets were encoded as pseudowords; therefore priming based on episodic retrieval should be facilitating (Zeelenberg et al, 2004). D-Target pseudowords that had been presented as distractors only on the preceding day were not necessarily processed and encoded as specific nonword entities, but rather appear just familiar and more word-like. This should lead to inhibitory effects on the next day. Further, if D-target pseudowords develop lexical characteristics, lexical decisions on neighbour words should be inhibited. On the other hand, if pseudoword repetition priming is purely based on the activation of neighbour word representation as in Experiment 5, then presentation of pseudowords should facilitate word lexical decisions on the next day.

7.1.1 Methods and Materials

Participants. Sixteen students (one male, age 19 – 32 years, mean 22.6 years) from Technical University Braunschweig were tested in three one hour sessions with Session 1 and Session 2 at least 2 days apart, and Session 2 and Session 3 exactly one day apart. All were native German speakers and had normal or corrected to normal vision. They took part for course credit.

Task. In the first two sessions (training and study session) participants passively watched word streams of 42 to 58 items, and made a lexical decision on the last item of each word stream. To reduce conscious processing and encoding of each item in episodic memory, no task had to be performed during the word streams. On Session 3 (experimental session) half of the participants performed lexical decisions on singly presented words and pseudowords, the other half had to report whether they had seen the singly presented items in the study session one day before. No primes were presented in Session 3.

Stimuli. Three non-overlapping stimulus sets were used: the target set, the distractor set, and the filler set. For the *target set* 160 LF-words and 160 HF-words were chosen from WebCelex (Max Planck institute for Psycholinguistics, 2001). Each target word was assigned one word neighbour, one pseudoword neighbour and one unrelated word. Word neighbours differed from target words in one letter, were of the same frequency class as the target word and did not have any other neighbours in the stimulus set. Pseudoword neighbours were nonwords that differed from the target word in exactly the same letter as the word neighbour. The replacing letter was chosen randomly so that the pseudoword still was legal in regard to German grammar and pronunciation. Unrelated words were chosen so that the overlap with the target word was at maximum 2 letters. Another 160 LF-words and 160 HF-words served as word neighbour primes for pseudoword targets. In addition, each of these pseudoword targets was assigned one pseudoword neighbour and on unrelated pseudoword as primes. Again pseudoword neighbours and word neighbours differed in the same letter position from the target pseudoword. The *distractor set* were comprised of 384 words and pseudowords, which occurred in word streams only without being a target for lexical decision. Pseudoword distractors were neighbours from exactly one word distractor. For every participant only the pseudoword or the word distractor occurred in the whole experiment, but not both. The filler set another 100 words and pseudowords each served as filler items, which were neither lexical decision targets nor recognition targets. Distractor and Filler items did not have any neighbour word or pseudoword within the prime-target set (see table 9 for characteristics of Target set and Table 10 for characteristics of Distractors and Fillers).

Table 9: Stimuli characteristics for the target set of experiment 5: means (SD) for item length, number of syllables, frequency counts, overlap, and means (SD) / medians for letter change position. w: words; pw: pseudowords

		Letters mean (SD)	Syllables mean (SD)	CELEX frequency count/million	CELEX frequency log/million	Trigram frequency	Change position	Overlap
LF-w	Targets	5.0 (1.0)	1.7 (0.5)	3.2 (1.8)	0.43 (0.27)	163.8 (19.4)	--	--
	Word neighbour	5.0 (1.0)	1.7 (0.5)	3.1 (1.7)	0.40 (0.28)	164.1 (19.9)	2.1 (1.4) 1	0.81 (0.05)
	Pseudo-word neighbour	5.0 (1.0)	1.7 (0.5)	--	--	161.6 (19.5)		0.81 (0.06)
	Unrelated	5.0 (1.0)	1.7 (0.5)	2.8 (1.5)	0.37 (0.26)	151.0 (24.8)	--	0.09 (0.12)
HF-w	Targets	5.0 (1.0)	1.5 (0.5)	81.2 (148.3)	1.60 (0.48)	170.9 (16.1)	--	--
	Word neighbour	5.0 (1.0)	1.5 (0.5)	52.8 (68.9)	1.48 (0.43)	170.9 (16.9)	2.3 (1.5) 2	0.81 (0.06)
	Pseudo-word neighbour	5.0 (1.0)	1.5 (0.5)	--	--	165.5 (19.1)		0.81 (0.06)
	Unrelated	5.0 (1.0)	1.5 (0.5)	60.5 (102.6)	1.50 (0.44)	155.3 (24.0)	--	0.08 (0.13)
LF-pw	Targets	5.0 (1.0)	1.7 (0.5)	--	--	154.2 (24.9)	--	--
	Word neighbour	5.0 (1.0)	1.7 (0.5)	2.6 (1.7)	0.32 (0.29)	156.4 (25.1)	2.1 (1.4) 1	0.82 (0.06)
	Pseudo-word neighbour	5.0 (1.0)	1.7 (0.5)	--	--	155.0 (24.3)		0.82 (0.06)
	Unrelated	5.0 (1.0)	1.7 (0.5)	--	--	148.2 (25.8)	--	0.24 (0.18)
HF-pw	Targets	5.0 (1.0)	1.7 (0.5)	--	--	154.6 (25.9)		--
	Word neighbour	5.0 (1.0)	1.7 (0.5)	93.7 (218.1)	1.60 (0.50)	156.9 (27.0)	2.3 (1.4) 2	0.81 (0.06)
	Pseudo-word neighbour	5.0 (1.0)	1.7 (0.5)	--	--	155.0 (25.8)		0.82 (0.06)
	Unrelated	5.0 (1.0)	1.7 (0.5)	--	--	151.8 (22.9)	--	0.25 (0.20)

Table 10: Stimuli characteristics for distractor set (a) and filler set (b) of Experiment 6

a)	Letters	Syllables	CELEX frequency		Trigram frequency	Change posititon
			count/million	log/million		
LF-words	5.5 (0.7)	1.7 (0.5)	3.7 (2.7)	0.42 (0.27)	159.4 (24.2)	--
HF-words	5.6 (0.6)	1.7 (0.5)	46.9 (79.0)	1.36 (0.47)	163.8 (22.7)	--
LF-pseudowords	5.5 (0.7)	1.7 (0.5)	--	--	153.1 (22.6)	3.3 (1.5) 3
HF-pseudowords	5.6 (0.6)	1.7 (0.5)	--	--	152.5 (25.5)	3.4 (1.6) 3
b)						
Words	5.6 (0.9)	2.1 (0.6)	21.9 (33.6)	0.90 (0.64)	153.4 (21.8)	--
Pseudowords			-- --	-- --	152.7 (19.9)	2.9 (1.6) 3

Design. Session 1 and 2 consisted of 320 trials in 8 blocks, each (i.e., 40 trials per block). Each trial consisted of a short word stream (42-58 items) and one lexical decision immediately following the last item in a stream. A word stream contained one prime together with distractors and filler items. Primes either were repetition primes, word neighbours, pseudoword neighbours or (semantically and orthographically) unrelated primes, and were presented in the word stream either once or five times. Distractor items were also presented once or five times. Overall 3 words and 3 pseudowords were presented 5 times in each word stream, 12 to 28 items were presented once. The average spacing between two repetitions was 8 items plus/minus 2. Lexical decision targets were low frequency words, high frequency words, low frequency pseudowords or high frequency pseudowords (targets form the target set). Each target occurred only once in the study sessions, distractors occurred repeatedly in different word streams between 30 and 40 times in one session. These independent variables (Prime Type, Presentation Frequency, Target Type) built a full factorial design of 32 different conditions with 20 trials per condition. The allocation of targets to conditions for each item type was counterbalanced over participants. Trials were randomly mixed across the whole experiment.

The experimental session consisted of 1024 trials in 16 blocks (i.e., 64 trials per block). Targets were chosen from the target and distractor set individually for each participant: First, all items that had occurred in Session 2 as targets were assigned T-targets. Second, unrelated primes from the target set that did not occur in Session 1 and Session 2 served as control targets, because these were matched with respect to word length and frequency. Third, words and pseudowords from the distractor pool that only occurred in word streams in Session 2 but never as lexical decision targets, were assigned D-targets. Fourth, words and pseudowords from the distractor pool that had not occurred before served as control for the distractor targets. Fifth, neighbours of words and pseudowords that had occurred as distractors were assigned neighbour targets, e.g. if the word “Wurzel” had occurred as distractor in the study session, the pseudoword “Wurwel” was assigned a neighbour target in the experimental session. Main independent variables were target (old, new, neighbour), word type (LF-words, HF-words, LF-pseudowords, HF-pseudowords) and target type (target, distractor). Trials from all experimental conditions were randomly intermingled.

Procedure. Figure 16 depicts the trial events of one trial in sessions 1 and 2. First, a fixation cross appeared for 1 second. Each word stream item was presented for 54 ms followed by a blank screen for 26 ms. Word stream length varied between 42 and 58 items. Immediately after the last item of the word stream the target was presented and remained on the screen until the participants responded by pressing the left or the right button for words and pseudowords, respectively. Response mapping was counterbalanced over participants. Inter-trial interval varied randomly between 1000 ms and 1400 ms.

In the experimental session, each trial began with a fixation cross centred on the screen for 1000 ms, followed by the target item, on which participants made a lexical decision or recognition response. On lexical decision trials targets were presented for a maximum of 2 seconds, trials with longer response times were scored as false. False responses were signalled by a short tone. On recognition trials, targets remained until the response and no feedback was given.

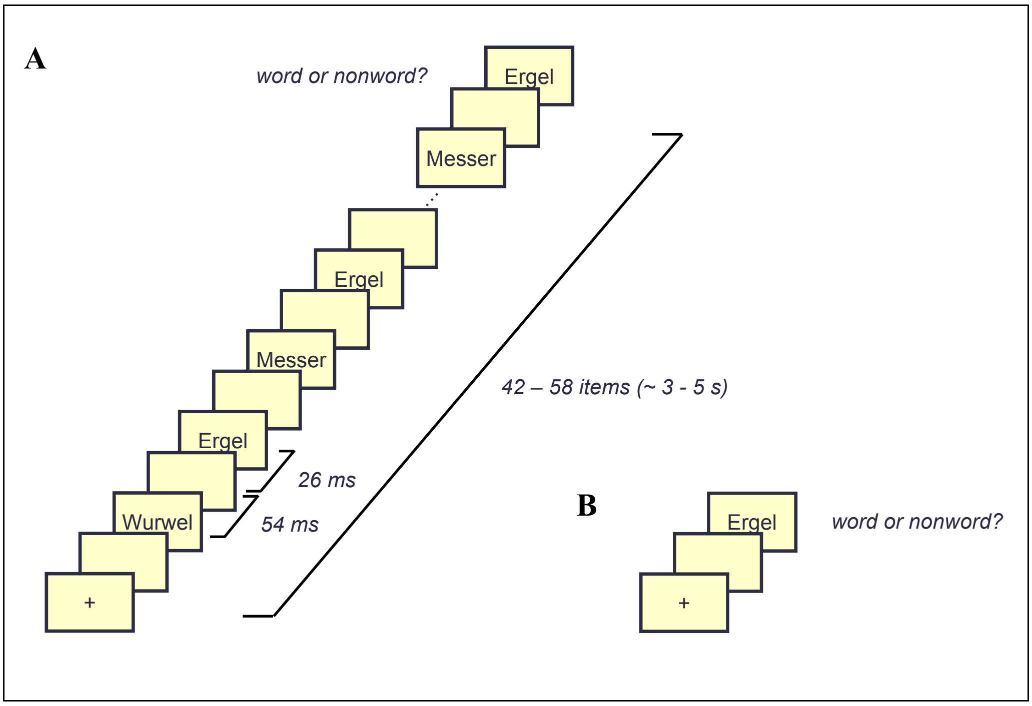

Figure 16: Trial events in training session (A) and experimental session (B) of Experiment 6a.

7.1.2 Results

Study sessions. Mean error rate for lexical decisions was 3.1 %. Response time data are not reported. Because of too few trials per condition and the target presentation immediately after the last word stream item, no priming effects at all were observable. Experiment 6b deals with this issue in more detail.

Lexical decision task. First, all incorrect responses were discarded from the data set (5.5%). Second, all T-target trials that had yielded incorrect responses in Session 2 were excluded (3.5% of all T-target trials). Third, 3.1% of all data were identified as outliers, and were excluded from all further analyses. In the end, 90.4% of all data were available for analysis. Figure 17 depicts mean response times for T-targets and D-targets. Responses on words were faster for HF-words than for LF-words [t = 13.04, p < 0.001]. Old words were responded to faster than new words [t = -4.10, p < 0.001], and T-targets were responded to faster than D-targets [t = -4.10, p < 0.001]. The main effects of novelty and target type showed a significant interaction [t = 4.49, p < 0.001]: whereas for T-targets responses on old words clearly were faster than on new words, response times for old and new words were virtually the same for D-targets. The interactions of Novelty X Word Type [t = 1.65, p=0.10], Word Type X Target Type [t = -0.07, p=0.95], and the Novelty X Word Type X Target Type [t = -0.17, p = 0.87] did not reach significance. Both random effects are substantial for the model fit, dropping Subject [χ^2 (1) = 56.64, p < 0.001] or Item [χ^2(1) = 2624.0, p < 0.001] both led to a worse fit. It is notable that for LF-words and HF-words, response times on new T-targets, new D-targets and old D-targets all are essentially the same (~640 ms for LF words, ~587 ms for HF words), only responses on old T-targets are clearly faster (608 and 565 for LF words and HF words, respectively).

A separate analysis conducted for D-targets including all three levels of Novelty (o*ld, new, and neighbour*) did not reveal any influence of repeated pseudoword presentation on lexical decisions to words [all |t|<= 1, p > 0.25], although Figure 17 (right panel) suggest an inhibitory effect on high frequency words.

Response times on pseudoword trials were not modulated by Pseudoword Type [t = 0.08, p = 0.94] or Target Type [t = -0.79, p = 0.44], but responses to old pseudowords were slower than responses to new pseudowords [t = 2.93, p = 0.003]. This effect is profound for D-targets but not for T-targets as shown by the significant interaction of Novelty X Target Type [t = 2.35, p = 0.02]. Further, the novelty effect for pseudowords is more profound for HF-pseudowords than for LF-pseudowords [t = 2.56, p = 0.01]. No interaction of Pseudoword Type X Target Type [t = 0.58, p = 0.56] or Pseudoword Type X Target Type X Novelty [t = 0.77, p = 0.44] was observed. The LME model containing random factors Subject and Item, outperform both LME models dropping Item [χ^2 (1) = 34.95, p < 0.001] or dropping Subject [χ^2 (1) = 3239.50, p < 0.001]. Note that response times for T-targets and new D-targets are in a small range from ~630 ms to 640 ms, and only responses to old D-targets have a increased mean response time of ~660ms.

A separate analysis for D-targets including all three levels of Novelty (*old, new, neighbour*) revealed an inhibitory influence of neighbour primes on lexical decisions of pseudowords [t = -2.07, p = 0.04].

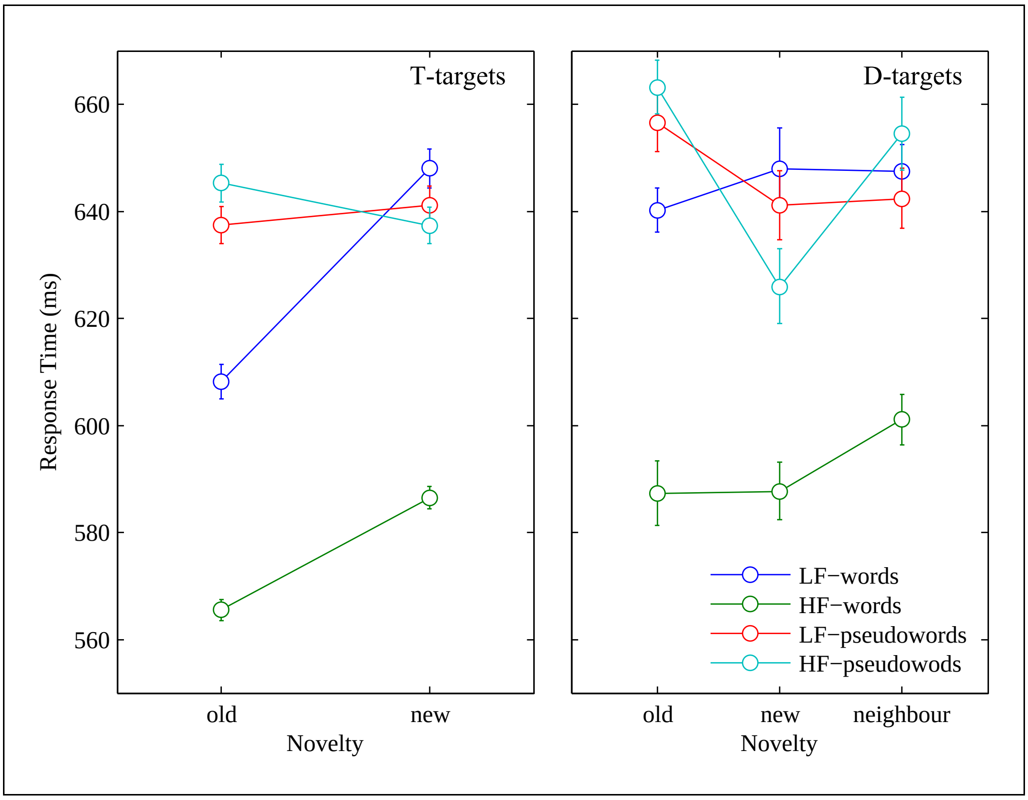

Figure 17: Mean lexical decision response times on Session 2 in Experiment 6a. For targets from Session 1 there is a facilitation effect for words, but not for pseudowords (left panel); this pattern is reversed for distractors from Session 1 (right panel). Error bars represent standard error of mean.

Recognition task. 7.2% of old T-targets, which were classified incorrectly in the lexical decision task in Session 2, were excluded (2.6% of all data). Figure 18 shows the mean percentage of 'old'-responses for old items (correct recognition) and new items (false alarms). Table 11 shows recognition performance in terms of d'. A LME model with a binomial link function was fitted to the binary recognition data. For words, recognition performance is substantially better for T-targets than for D-targets [z = 7.2, p < 0.001], and new words are better recognized than old words [z = -14.5, p < 0.001]. Recognition for new words is virtually identical for T-target words and D-target words, but old T-target words are better recognized than old D-target words [z = -6.9, p < 0.001]. LF-words are slightly better recognized than HF-words, but this main effect does not reach significance [z = 1.4, p = 0.16, all other |z|<1, p > 0.33].

For pseudowords, recognition was slightly better for T-targets than for D-targets, but did not reach significance [z = 1.9, p = 0.06], and new pseudowords were better recognized than old pseudowords [z = -29.2, p < 0.001]. Again, recognition of new pseudowords is similar for T-targets and D-targets, but recognition of old T-targets was better than recognition of old D-targets [$z = -5.1$, $p < 0.001$]. Figure 18 reveals a slight trend to better recognition for HF-pseudowords, which is more profound for T-targets, but neither the main effect of frequency [$z = -1.8$, $p = 0.07$] nor the interaction Frequency X Novelty [$z = -1.4$, $p = 0.17$] were statistical reliable. No other term was significant [all other $|z|<1$, $p > 0.7$]. Table 11 shows individual recognition performance d'. Note, that performance for D-targets is worse than for T-targets in each single case. To evaluate if recognition exceeds chance level, one-sided t-tests on d' were performed for each condition. Performance on T-targets clearly exceeds chance level for all item types, but mean d' on D-targets are quite small and only for HF-pseudowords significant different from chance (Table 11).

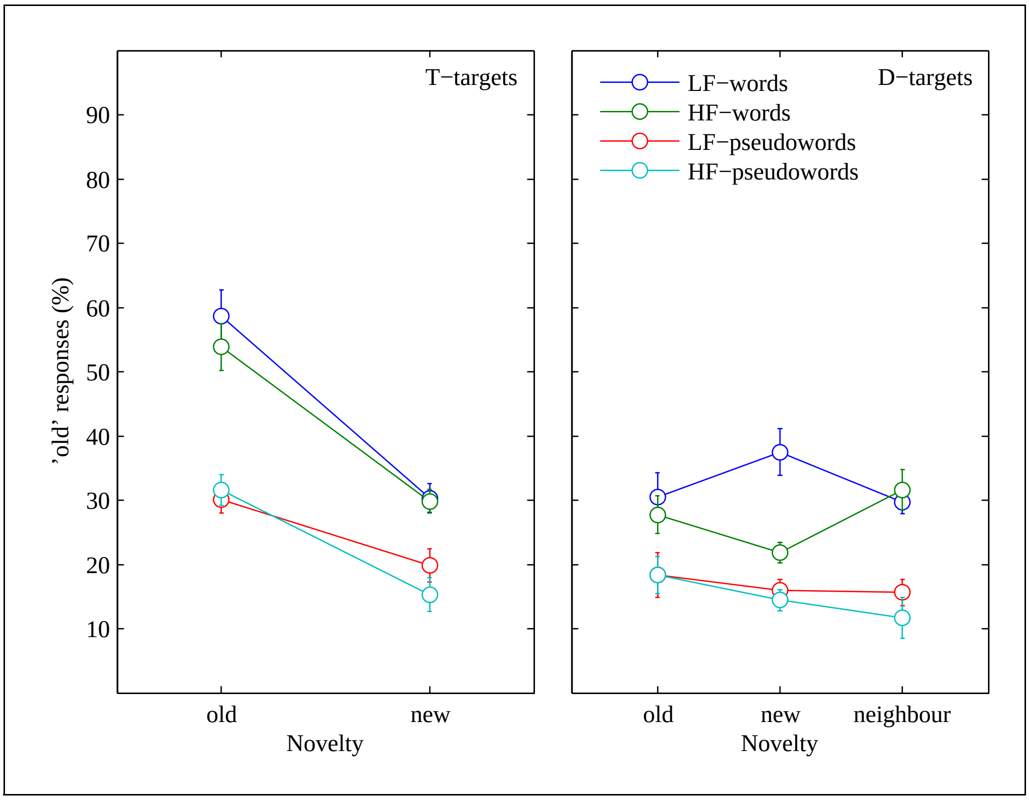

Figure 18: Recognition performance in Experiment 6a. Error bars represent standard error of mean.

Table 11: Recognition performance (d') for each participant and target

		1	2	3	4	5	6	7	8	mean	t	p-value
LF-words	T-targets	0,89	1,03	0,94	0,85	0,79	0,36	0,67	1,01	0,82	10.6	<0.001
	D-targets	0,21	-0,21	0,24	0,00	-0,26	0,00	-0,54	0,65	0,01	0.1	0.47
HF-words	T-targets	1,14	1,04	0,59	0,35	0,68	0,54	0,64	0,48	0,68	7.1	<0.001
	D-targets	-0,86	0,17	-0,16	0,00	-0,24	0,00	-0,08	0,08	-0,14	-1.2	0.86
LF-pseudo-words	T-targets	0,74	0,26	0,18	0,53	0,60	0,09	0,25	0,63	0,41	4.8	0.001
	D-targets	0,85	0,12	-0,52	0,62	0,40	-0,20	-0,10	0,21	0,19	1.1	0.14
HF-pseudo-words	T-targets	0,73	0,73	0,49	1,01	0,86	0,34	0,56	0,87	0,68	8.9	<0.001
	D-targets	0,22	-0,12	0,14	0,33	-0,26	0,33	0,61	0,80	0,29	1.9	0.05

7.1.3 Discussion

In lexical decisions substantial repetition priming was observed for words that had been lexical decision targets one day before, but not for words that had been distractors within word streams. In contrast, lexical decisions on pseudowords were inhibited when pseudowords had been presented in word streams only, but were not affected when pseudowords had been lexical decision targets one day before. These repetition effects are dissociated from recognition performance in that words and pseudowords were better recognized when they served as lexical decision targets one day before, than when they had been distractors one day before. Recognition performance for the latter hardly exceeded chance level. Thus, repetition effects for T-targets can be explained by an episodic component or response learning, whereas the effects found for D-targets seem not to be based on episodic memory retrieval.

Priming for pseudowords that had been targets before should have been facilitatory when based on episodic retrieval, because the episode of the last session includes the nonword status of each pseudoword. However, the recognition rate of pseudowords was relative low, i.e., episodic traces can be recalled only for a few pseudowords. Therefore both positive (episodic based) effects and negative (lexical based) effects may be mixed in the experiment leading to an overall null effect.

Neighbour primes showed unclear results. Pseudowords, which were repeated within word streams, showed a trend to inhibit lexical decisions on high frequency words. This can be interpreted as evidence for a growing word-likeness of pseudowords due to multiple presentations, and the development of lexical characteristics as found by Bowers et al. (2005) and Gaskell and Dumay (2003). On the other hand, repeated presentation of words caused lexical decisions on pseudowords to be inhibited, although multiple presentation of words did not affect lexical decisions on the same words. Thus, the presentation of words leads to lasting changes in the memory system that become visible only on pseudoword trials. It may be that the changes are too small to affect the word response on the next day, but then a small trend should be visible. It is more plausible that these finding indicates different mechanisms of word and pseudoword decisions.

7.2 Experiment 6b

In the study sessions of Experiment 6a, no priming was observed, although the word streams were much shorter than in the preceding experiments. One possible explanation is that the presentation of the lexical decision target immediately after the last word stream item impeded priming effects. Because participants did not receive a cue, when to expect the target, they could recognize a target as target not until its duration exceeded the duration of “normal” word stream items. So, Participants had to make two decisions: Is the item presented at the moment a target? If so, is it a word or a pseudoword?

Therefore, I modified the setup by inserting a task between word stream and lexical decision task. The beginning of the lexical decision was cued, so that participants knew when to concentrate on lexical decision. With this modified version I expected stable repetition priming effects larger than in Experiments 1 to 5.

7.2.1 Methods and Materials

Participants. Eleven students (four male, age 20 – 37 years, mean 24.6 years) from Technical University Braunschweig were tested in three one hour sessions with at least 2 days apart. All were native German speakers and had normal or corrected to normal vision. They took part for course credit. Due to partial data loss, one data from one participant was excluded from all analyses.

Task. Participants were instructed to monitor a stream of rapidly presented words and pseudowords. Immediately after, they performed a 2-AFC recognition decision between two words or two pseudowords, followed by a lexical decision on a different word or pseudoword.

Stimuli. Word streams were constructed the same way as in Experiment 6a except that repetition and unrelated primes occurred only. Each repetition prime could occur once or five times within a word stream. One third of all trials included no repetition prime, one third included one prime presentation, and one third included five prime presentations. In each trial only one prime from the target set occurred, all other items originated from the distractor set and the filler set. In each word stream three pseudowords and three words were presented five times, and six to fourteen words and pseudowords were presented once (depending on the word stream length). Items presented five times occurred once in every quintile of the word stream, items presented once occurred in one of the quintiles. Quintile number was counterbalanced across targets for each participant. The spacing between two occurrences of one item was set to eight items on average with a jitter of plus/minus two items.

Design. The experiment consisted of 640 trials divided into three sessions with six blocks of 35-36 trials each. Independent variables were Word Type (LF-words, HF-words, LF-pseudowords, HF-pseudowords) and Presentation Frequency (0, 1, 5). All independent variables were manipulated within participants. For each Word Type two repetition conditions contained 53 trials and one contained 54 trials counterbalanced over participants.

Procedure. Trials started with a fixation cross for 1000 ms, followed by a blank screen for 500 ms. Then the stream presentation began with a speed of 54 ms per item followed by a blank screen for 26 ms. In the first three and the last three positions of a word stream a string of 8 symbols was presented (‘########’, ’%%%%%%%%’, ’&&&&&&&&’). Immediately after the last symbol string at the end of a stream the forced choice recognition task was presented: A fixation cross was centred on the screen and 3.2° of visual angle above and below this fixation cross, two alternative words or pseudowords were presented, one of which had occurred either once or five times in the preceding stream. Participants responded with the index finger (‘°-key’) and thumb of the left hand (left shift key) whether they had seen the upper or the lower alternative in the word stream. No feedback was given. After an interval of 1000 ms to 1400 ms, a red fixation box cued the following lexical decision for 1000 ms followed by a blank screen for 250 ms. Lexical decision targets were presented centred on the screen and remained until response. Participants responded by pressing the left arrow or right arrow key with the index finger or middle finger of the right hand, respectively. The key-response mapping was counterbalanced across participants. Incorrect responses were indicated by an acoustical signal. Inter-trial interval was set to 1500 ms. Figure 19 depicts the order of trial events. Before the start of each block an instruction screen was presented, which showed the block number and reminded participants of the key-response mapping (Figure 20).

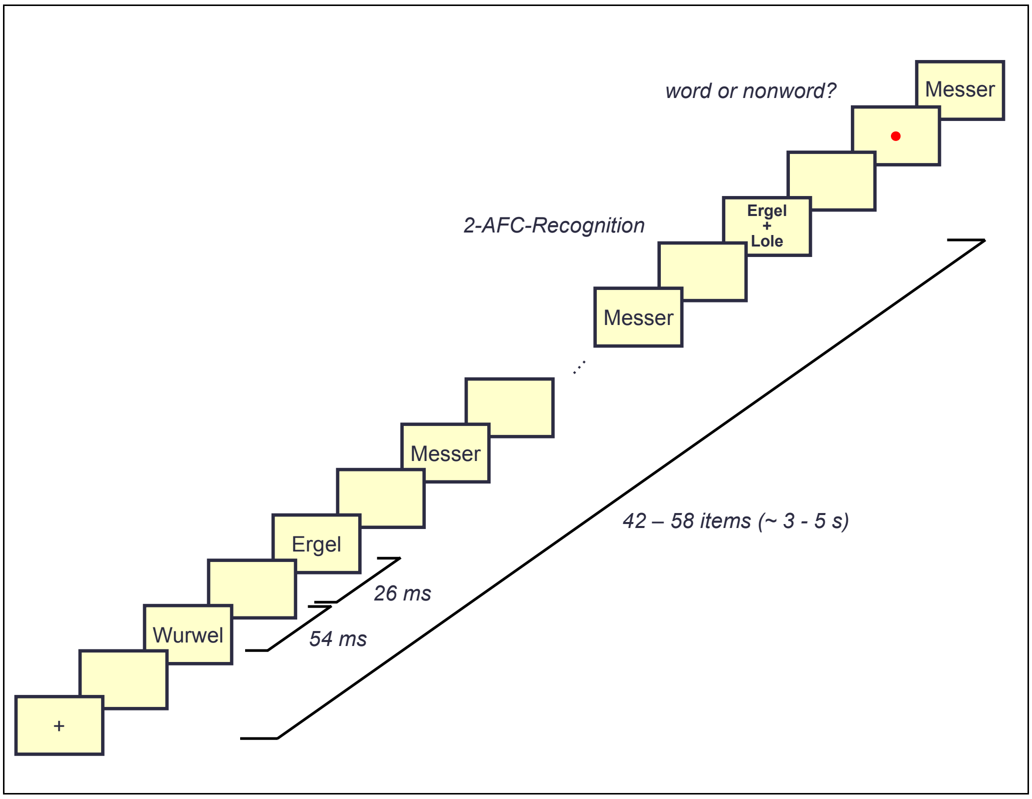

Figure 19: Trial events in Experiment 6b

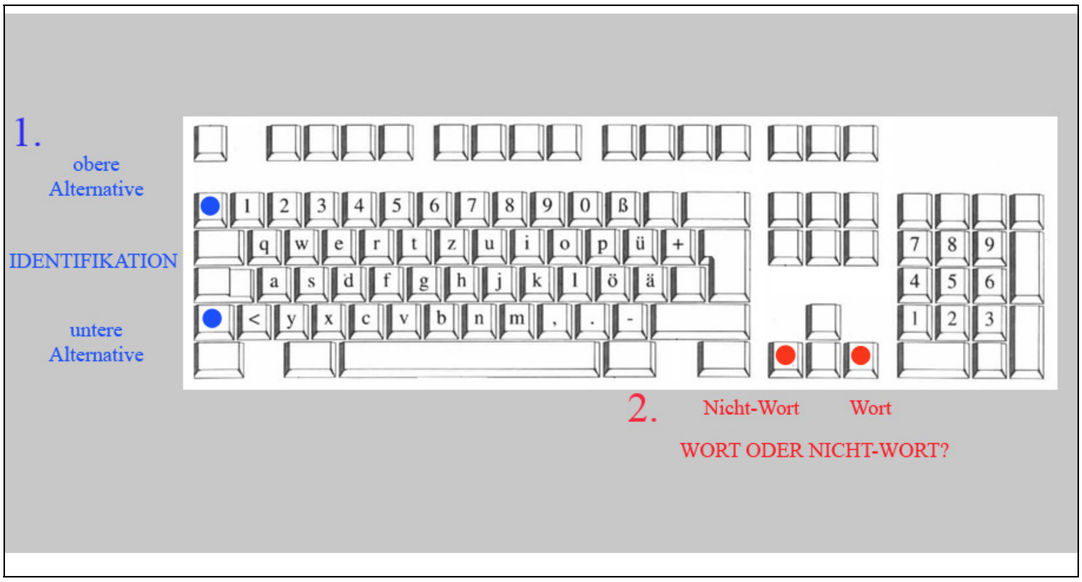

Figure 20: Instruction Screen used in Experiment 6. Response keys for the recognition task are shown in blue (upper key for upper alternative, lower key for lower alternative, response keys for lexical decisions are shown in red. On the keyboard the appropriate keys were marked with the corresponding colours.

7.2.2 Results

Lexical decision task. 2.1 % of all data were excluded as outliers. Mean response times are shown in Figure 21 (left panel). Responses were faster to HF-words than to LF-words [583 ms vs. 639 ms, t = 9.63, p < 0.001] and response times decreased with increasing Presentation Frequency, as confirmed by a reliable linear trend [t = 5.2, p < 0.001]. This decrease seemed more profound for LF-words, but the interaction term did not reach significance [t = -1.24, p = 0.22]. Dropping the random terms Subject or Item always reduced the goodness-of-fit [χ^2 (1) = 1181.6, p < 0.001 and χ^2 (1) = 64.8, p < 0.001].

Response times to pseudowords increased with Presentation Frequency [t = -3.33, p = 0.002], but no main effect for pseudoword frequency [t = 1.24, p = 0.21] and no reliable interaction were observed [t = -0.51, p = 0.62]. Again, dropping one of the random terms worsened the model fit [χ^2 (1) = 1799.4, p < 0.001 and χ^2 (1) = 100.5, p < 0.001].

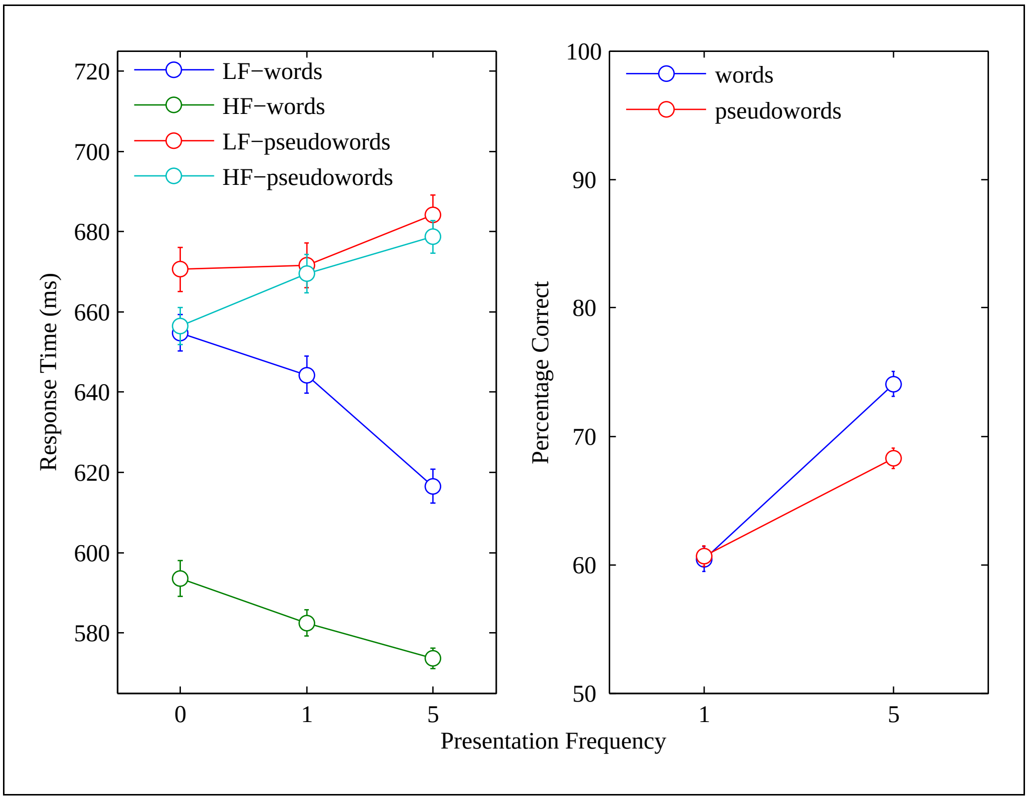

Figure 21: Mean response times of lexical decisions (left panel) and mean recognition performance in 2-AFC-task (right panel) for Experiment 6b. Error bars represent standard error of mean.

Recognition task. As can be seen from Figure 21 (right panel) recognition performance was well above chance, and increased from one presentation to five presentations for both words and pseudowords. A logistic regression LME model yielded significant effects for words and pseudowords [z = -7.82, p < 0.001 and z = -4.27, p < 0.001, respectively]. To investigate the relationship of recognition performance and priming in lexical decision, a LME model with Presentation Frequency, Word Type and Recognition Performance (correct vs. incorrect) as fixed factors was fitted to the data. Because the introduction of the factor Recognition resulted in a highly unbalanced design, significance was evaluated by model comparison with likelihood ratio statistics. Neither inclusion of a main effect for recognition [χ^2 (1) = 0.14, p = 0.71] nor inclusion of any second order interaction [χ^2 (2) = 2.7767, p= 0.25 and χ^2 (3) = 1.113, p = 0.774] or third order interaction [χ^2 (6) = 5.2, p = 0.52] improves the model fit compared to the standard model.

7.2.3 Discussion

In this modified "word shower" paradigm with shortened word streams and only one target per word stream, substantial cumulative priming effects were observed. Word recognition was clearly above chance and was better for items presented five times than for items presented once. As in the preceding experiments, priming effects were not modulated by word frequency, which highlights the central role of processing time for frequency modulation (Versace & Nevers, 2003). In comparison with the results of Experiment 3 priming effects were slightly larger as was recognition performance, but again recognition was unrelated to lexical decision. Thus, conscious access to an encoding episode of an item does not modulate the priming effect this encoding episode has on subsequent processing of the same stimulus. This corroborates findings from Ratcliff et al. (1985). However, it has to be noted that recognition and priming were assessed in the same trial, but on different items. Therefore, an incorrect response on a recognition target does not necessarily imply missing recognition of the lexical decision item. Rather, recognition performance can be taken as a global measure of the processing of individual items in a word stream.

The results allow inferences about conscious awareness of items presented in rapid order. In Experiments 1 to 5, recognition performance was near chance, but did not allow conclusions about whether items never had been conscious or had just been "forgotten" until the time of test. In this experiment, the delay between presentation of an item and recognition test was considerably shorter. If items presented in word streams never had reached consciousness, recognition performance should have been near zero. But performance in all conditions and for all item types was well above chance and increased with presentation frequency.

So, what are the advantages and disadvantages of employing long streams and short streams? First, short word streams allow for better control of interfering variables, because only one critical item is included in a stream and only few distractors may interfere with the processing of the critical item. On the other hand, long word streams are more efficient in that a smaller number of word streams and less distractors are needed to induce priming in a great number of words. Second, the recognition rate obtained with short word streams is much higher than with long word streams and differs to a greater amount between different presentation frequencies. Thus, long word streams of several hundred items can be used to induce repetition priming while recognition remains near chance, whereas in short word streams recognition performance varies between different experimental conditions. Amazingly, the amount of priming in short streams seems only slightly elevated as compared to long streams: For words, priming was approximately 30 ms in Experiment 6b (short streams) and approximately 20 ms in Experiment 3 (long streams). Therefore long word streams are recommended for the investigation of “pure” repetition priming (without contamination by conscious recognition), short word streams are more useful to investigate effects of distractors on individual items, and in cases where the recognition rate is not important.

8. General discussion

The present experiments gave clear evidence for cumulative repetition priming of words and pseudowords presented in a rapid serial visual stream of more than 200 items. The "word shower" paradigm is suitable to induce long-term priming effects, although the processing of each single item is severely degraded, and cannot be recollected consciously at the time of test. The cumulative nature of priming was replicated in six experiments, all of which showed that repetition effects increased with increasing presentation frequency.

Massive repetition priming for words and pseudowords. Words consistently showed facilitatory priming, whereas pseudowords consistently showed inhibitory priming. Repetition priming of pseudowords was in general smaller than that of words. This is what both MROM and REM-LD predict: On presentation of words, evidence for specific words is accumulated and the lexical decision on a later test is facilitated. On presentation of pseudowords, evidence for the neighbour word is accumulated and the lexical decision is inhibited. Because the 'word' evidence on pseudoword trials is less than on word trials, priming should be less in magnitude.

Besides the cumulative priming effects within word streams, Experiment 1a and 1b also provided evidence for cumulative repetition priming of words with primes presented once in every word stream. Moreover, the effects seemed to be at least as high in amount as for repetitions within word streams. But as discussed in Experiment 1a, the amount of priming in the two different conditions cannot be compared because of different items and especially because of a different number of items in each condition. Nevertheless, the complete absence of priming for pseudowords repeated across word streams is puzzling, especially in light of the long lasting effects when repeated within word streams (Experiment 4). However, this may point to different mechanisms for repetition priming of words and pseudowords as assumed e.g. by Grainger and Jacobs (1992, 1996).

Word frequency effects. In long-term priming paradigms typically low frequency words benefit more from repetition than high frequency words. Forster and Davis (1984) have argued that this frequency effect occurs only if the first presentation already requires active processing and responding, but not when it is viewed passively. Versace and Nevers (2003) demonstrated that not the task demands but rather prime duration is the main influencing variable. To obtain word frequency effects, the prime has to be processed for a certain time. If this *minimal necessary processing time* is not available, words benefit from repetition regardless of their frequency. This finding is corroborated by the present experiments: A clear interaction was observed in Experiment 1a only, where the effective processing time was 250 ms for each item. In Experiment 2 a low frequency advantage was observed for items presented once only, which had an effective processing time of more than half a second. In all other experiments the effective processing time was reduced to 80 ms, and no reliable advantage for low frequency words could be observed. Therefore the *minimal necessary processing time*, postulated by Versace and Nevers (2003) has to lie somewhere between 80 ms and 250 ms.

Pseudoword frequency effects. Experiments 3 to 6 examined differences between pseudowords based on the frequency of their parent words, i.e., are there differences in priming between pseudowords that had been created from low frequency words as compared to pseudowords that had been created from high frequency words. Both MROM and REM-LD predict faster responses to low frequency pseudowords than to high frequency pseudowords, and Perea et al. (2005) reported a small response time advantage of low frequency pseudowords. So it is of interest if these differences can be replicated and differential priming effects can be obtained. Experiment 3 yielded slightly faster responses for low frequency pseudowords, but in Experiment 4 responses to low frequency pseudowords were slower than to high frequency pseudowords, and Experiments 5 and 6b did not yield any differences in response times for low frequency and high frequency pseudowords. Further, no differences in priming were obtained in any experiment. Thus, the present experiments do not yield evidence for frequency differences of pseudowords. However, it may be that the stimuli used in the present experiments were not appropriately chosen to examine such small effects. Pseudowords were created from words of different frequency classes, but it was not controlled whether a pseudoword had a neighbour word that was not part of the stimulus set. So it may be, for example, that a low frequency word as “NARR” was transformed to the low frequency pseudoword “NART”. But “NART” could also have been created from the high frequency word “BART”. Therefore, it can be considered a high frequency pseudoword, too. Thus, a more carefully chosen stimulus set or a different variable of “pseudoword frequency” would have been more conclusive.

Impact of conscious recollection. An important question is whether the observed repetition effects are based on the accumulation of word evidence, or whether they in fact reflect the higher probability of conscious detection and recollection of items presented more frequently. In Experiment 2 I manipulated the probability of conscious recollection by decreasing the presentation duration with increasing presentation frequency. The results clearly demonstrated repetition effects, although the probability of detecting an item in a word stream was higher for items presented once. Experiment 3 quantified further the relation between recognition and priming on an individual item level, when the individual presentation duration of each item was equal. Recognition performance increased with more presentations, but repetition effects were still observed when differences in recognition performance were statistically controlled. Although recognition performance was above chance in several experiments, massive repetition priming seems not to depend on conscious recollection at test, because cumulative priming effects were still observable when recognition was equal for different presentation frequencies.

Longevity of priming effects. Experiments 1a and 1b demonstrated that cumulative priming effects are independent of the position in the word stream as well as the trial number of test, i.e., they proved to be independent of the study-test delay (Experiments 1a and 1b). These findings indicate a long-lasting effect that is not prone to rapid decay, although the effective processing time of each prime was only 80 ms. Versace and Nevers (2003) reported that effects for very briefly presented primes decay in up to 3000 ms. But the effects they found at a delay of 3000 ms still were approximately 15 ms in amount, which is in the same range observed in the present experiments. Therefore, the present findings do not contradict Versace and Nevers (2003), but corroborate and extend their findings in that a small priming effect remains even at long lags.

Moreover, Experiment 4 showed that priming does not decay, even when measured after seven minutes (i.e., after the following word stream). These findings clearly demonstrate the longevity of priming, and are in line with previous findings of a long-term repetition component (McKone, 1995; Ratcliff et al., 1985). Note that in contrast to these studies participants performed a lexical decision only in the test phase, whereas the multiple prime presentations were task irrelevant as in short-term priming paradigms. Thus, regardless of response selection the presentation of a word led to accumulation of evidence for the presented word. In terms of an activation approach the activation level of specific word forms decreases with a very slow rate.

In Experiment 6a, I investigated delays of one day. The findings highlight the importance of response learning, and further show a clear dissociation of word and pseudoword repetition priming. Words that had been lexical decision targets (response relevant) one day before showed clear facilitatory priming, whereas pseudowords that had been lexical decision targets the day before did not show any effect. In contrast, words that had been word stream distractors (i.e., not response relevant) the day before did not benefit from repetition, but pseudowords that had been distractors showed inhibitory priming effects. Amazingly, a word that had been a distractor inhibited the lexical decision to its pseudoword neighbour on the next day, although it did not facilitate the lexical decision to itself on the next day. Vice versa, there was a trend that massive repetition of pseudowords inhibits lexical decisions to high frequency words on the next day. This is in line with several findings on acquisition of new words (Bowers et al, 2005; Gaskell & Dumay, 2003). Further, these findings clearly indicate different priming mechanisms when primes are response relevant as compared to response non-relevant primes. If a prime is response relevant, for example, when it is a target for a lexical decision, then it is processed in a completely different way than primes presented briefly among others without instruction to process them. First, because a decision is made to a lexical decision target, it is processed more deeply than a distractor prime. Second, responding to a target probably involves the creation of a strong episodic trace, which contains not only the prime, but also the context, the response chosen by the participant, and the feedback, if the response was correct or not. A distractor prime only contains degraded information of the stimulus, but no response and no feedback, and possibly no episodic trace at all is formed, as shown by the recognition data of Experiment 6a: Recognition of target primes was above chance, but recognition of distractor primes was at chance.

Cross-priming of words and pseudowords. Experiment 5 focussed on the question whether the presentation of pseudowords activate the word form representation of their neighbour words as postulated by several models of word recognition and lexical decision (e.g. MROM, REM-LD). If so, words presented in the word stream should lead to (enhanced) inhibitory priming on pseudoword lexical decision, and pseudoword presentation in the word stream should lead to (reduced) facilitatory priming effect in word lexical decision. In fact, this data pattern was observed in Experiment 5.

In sum, the present findings indicate small but reliable cumulation of priming effects despite the presence of many distracting items in a word stream. They further corroborate previous findings that repetition priming is not a uniform phenomenon; rather, repetition effects can occur at different levels and may depend on different mechanisms. They also show that repetition effects are automatic in that minimal processing only is necessary to produce priming.

Lexical or episodic effects? MROM as a lexical model of word recognition and REM-LD as an episodic account of lexical decision and repetition priming both can explain most of the presented results. The fact that repetition priming across word streams leads to equal or even more priming (Experiments 1a and 1b), seems to favour an episodic explanation, because cumulative priming based on activation should diminish with increasing lag between presentations. However, as mentioned above, these results may depend on methodological issues. On the other hand, REM-LD assumes the same process for 'word'-responses and 'nonword'-responses, so that it is difficult to explain dissociations between repetition priming for words and pseudowords as found in Experiments 1 (facilitatory priming for words repeated across streams, but null effects for pseudowords) and Experiment 6b (word primes led to inhibited responses to pseudowords, but did not affect responses to words). MROM explicitly assumes different mechanisms: 'word'-responses are based on an adjustable activation threshold, whereas 'nonword'-responses are based on an adjustable temporal deadline. Therefore, the results presented here can be explained better by a modified activation based account that assumes the cumulation of word form activation due to multiple word presentation and that operates with different thresholds for words and pseudowords.

However, proponents of episodic accounts strongly argue that activation based models cannot account for long-lasting priming effects, because sustained high activity of word form representations is not functional or biologically plausible. In fact, in light of the presented long-lasting results, it must be questioned whether the initial notion of residual activity in the lexical network is tenable and biologically plausible. A sustained higher level of activity means that activation in the mental lexicon prevails for a rather long time without any decay for many items simultaneously. This would be without any function and quite senseless from an energetic viewpoint. A more appropriate term than *activation* would be *excitability.* Thus, Figure 1 (p. 27) does not show the *activity level* of word form representations, but rather an increasing *excitability* due to multiple word form representations. This would also be reasonable regarding neurobiological mechanisms of neural plasticity in the nervous system as e.g. synaptic facilitation or long-term potentiation, which are both based on an increasing supply of neurotransmitters due to synaptic activation in rapid succession. In turn, a greater supply of neurotransmitter leads to a greater excitability of a synapse (cf. e.g. Bliss & Collingridge, 1993; Citri & Malenka, 2007).

In a similar way cumulative priming effects may operate: On the first presentation the word form representation is activated and remains in an excitable state for a certain time. On the next presentation of the same word the word form reaches the same level of activation earlier, and its excitability is enhanced further. By this argumentation the fact that priming effects are still observable with repetitions across word streams becomes plausible: Because one presentation suffices to induce priming that holds at least until the next word stream, the excitability of the word form representation remains at a higher level too. Thus, presenting the prime word in the next word stream should lead to an enhanced effect. Episodic accounts like REM-LD lack any suggestions regarding neuronal mechanisms, e.g. how the addition of contextual features to an existing lexical trace is accomplished on the neural level.

The present results do not allow for a concluding statement whether the locus of effects is lexical or episodic in nature. A possible way to compare both accounts would be to investigate how the lag between two presentations influences massive repetition priming: Episodic accounts predict priming effects increasing with lag. With increasing lag the context differences between two presentations increase. The more different the context is at the second presentation, the more "new" information can be added to the episodic trace, which should lead to greater facilitation. In contrast, activation accounts predict decreasing effects, because the residual activity of the first presentation decreases as lag increases. Although a first step was taken in Experiments 1a and 1b by comparing repetitions within streams and across streams, no clear results were obtained.

One severe limitation of the presented experiments is that they consider lexical decisions only. It would have been valuable to investigate whether stable repetition priming occurs in tasks like naming, word completion, or a more semantic categorization. However, the major interest of the present study was to investigate repetition of nonwords, and, in particular, whether multiply presented nonwords become word-like. In research on identification of nonwords, consistently positive priming effects have been observed (e.g. Stark & McClelland, 2000; Bentin, 1989), and in such tasks the lexicality of nonwords is of no relevance. Thus, an interesting goal for further research to extend the "word shower" paradigm to other test tasks.

A second shortcoming may be the limitation to repetition priming, although a special kind of neighbourhood priming is dealt with in Experiment 5. Orthographic priming is of course an interesting question for the present paradigm, and it would be helpful in investigating the mechanisms of cumulative priming. However, the main purpose of the presented experiments was to develop a new paradigm for priming research, and to examine its possibilities and limitations. Repetition priming was chosen as the best suitable example, because a vast literature exists on different short-term and long-term priming paradigms, with different computational models and accounts.

The "word shower" proved to be a paradigm that is capable of inducing reliable long-lasting cumulative priming effects by very short but massive repetition. To this end it offers new possibilities for investigating the temporal dynamics of word recognition and implicit memory, and is therefore worthy further evaluation regarding other priming phenomena, other tasks and its neural correlates.

9. References

Baayen, R. H., Davidson, D. J., Bates, D. (submitted). Mixed-effects modeling with crossed random effects for subjects and items.

Baayen, R.H. (2007). The languageR package. R package version 0.2.

Baayen, R.H., Tweedie, F.J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon. *Brain and Language, 81*(1), 55-65.

Bates, D. M. (2007). lme4: Linear mixed-effects models using S4 classes. R package version 0.99875-7.

Bentin, S. (1989). Orthography and phonology in lexical decision: Evidence from repetition effects at different lags. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 61-72.

Bliss, T.V.P. & Collingridge, G.L. (1993). A synaptic model of memory: long-term potentiation in the hippocampus. *Nature,* 361, 31 – 39

Bodner, G. E. & Masson, M. E. J. (1997). Masked repetition priming of words and nonwords: Evidence for a nonlexical basis for priming. *Journal of Memory and Language, 37,* 268-293.

Bölte, J. (2001). Graded lexical activation by pseudowords in cross-modal semantic priming: Spreading of activation, backward priming, or repair? In J.D. Moore & K. Stenning (Hrsg.). *Proceedings der 23. jährlichen Konferenz der Cognitive Science Society* (pp. 86 - 91). Mawah, NJ: LEA.

Bowers, J. S. (2000a). In defense of abstractionist theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, 7, 83-99.

Bowers, J. S., Davis, C. J. & Hanley, D. A. (2005). Interfering neighbours: The impact of novel word learning on the identification of visually similar words. *Cognition, 97*, B45-B54.

Bowers, J.C. (2000b). The modality-specific and -nonspecific components of long-term priming are frequency sensitive. *Memory & Cognition*, 28, 406-414.

Bowers, J.S. (1994). Does implict memory extend to legal and illegal nonwords? *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 20, 534-549.

Brozinsky, C. J., Yonelinas, A. P., Kroll, N. E. A. & Ranganath, C. (2005). Lag-sensitive repetition suppression effects in the anterior parahippocampal gyrus. H*ippocampus, 15,* 557-561.

Cave, C. B. (1997). Very long-lasting priming in picture naming. *Psychological Science,* 8, 322-325.

Chun, M. M. & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Experimental Psychology: Human Perception and Performance, 21*(1), 109-127.

Citri, A & Malenka, R. C. (2007). Synaptic Plasticity: Multiple Forms, Functions, and Mechanisms. *Neuropsychopharmacology,* advance online publication, doi:10.1038/sj.npp.1301559

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335-359.

De Moor, W., Verguts, T. & Brysbaert, M. (2005). Testing the multiple in the multiple read-out model of visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1502-1508.

Deacon, D., Dynowska, A., Ritter, W. & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology, 41,* 60-74.

Dehaene, S., Naccache, L., Cohen, L., LeBihan, D., Mangin, J. F., Poline, J. B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4, 752-758.

Dobbins, I. G., Schnyer, D. M., Verfaellie, M., & Schacter, D. L. (2004). Cortical activity reductions during repetition priming can result from rapid response learning. *Nature, 428,* 316-319.

Evett, L.J. & Humphreys, G.W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology*, 33, 325-350.

Ferrand, L. (1996). The masked repetition priming effect dissipates when increasing the inter-stimulus interval: Evidence from word naming. *Acta Psychologica, 91,* 15-25.

Feustel, T.C., Shiffrin, R.M. & Salasoo, A. (1983). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General*, 112,309-346.

Forster, K.I. & Davic, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of experimental psychology: Learning, Memory, and Cognition*, 10, 680-698.

Friedrich, C. (2005). Neurophysiological correlates of mismatch in lexical access. *BMC Neuroscience, 6:64*

Gaskell, M. G. & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*, 105-132.

Graf, P. & Schacter, D.L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 501-518.

Grainger, J. & Jacobs, A.M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-Out Model. *Psychological Review*, 103, 518-565.

Greenwald, A. G., Draine, S. C. & Abrams, R. L. (1996). Three cognitive-markers of unconscious semantic activation. *Science,* 273, 1699-1702.

Grill-Spector, K., Henson, R. & Martin, A. (2006).Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*, 14-23.

Guillem, F., Rougier, A. & Claverie B. (1999). Short- and long-delay intracranial ERP repetition effects dissociate memory systems in the human brain. *Journal of Cognitive Neurosciences*, 11, 437-58.

Henson, R. N., Rylands, A., Ross, E., Vuilleumeir, P. & Rugg, M. D. (2004). The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *NeuroImage*, *21*, 1674-1689.

Henson, R.N., Shallice, T. & Dolan, R.J. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science, 287*, 1269-1272.

Holcomb P. J. & Grainger, J. (2006) On the time-course of visual word recognition: An ERP investigation using masked repetition priming. *Journal of Cognitive Neuroscience*,18,1632–1643.

Huber, D. E., Shiffrin, R. M., Lyle, K. B. & Ruys, K. I. (2001). Perception and preference in short-term word priming. *Psychological Review, 108*, 149-182.

Humphreys, G.W., Besner, D. & Quinlan, P.T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General*, 117, 51-67.

Hutchison, K. A., Neely, J. H., Neill W. T., & Walker, P. B. (2004). Is unconscious priming lexical or sublexical? *Consciousness and Cognition,* 13, 512-538.

Jacobs, A. M., Rey, A., Ziegler, J. C., & Grainger, J. (1998). MROM-P: An interactive activation, multiple read-out model of orthographic and phonological processes in visual word recognition. In J. Grainger, J. & A.M. Jacobs (Eds*.) Localist connectionist approaches to human cognition.*. Mahwah , NJ : Lawrence Erlbaum Associates.

Jacoby, L.L. (1983). Perceptual enhancement: persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 21-38.

Karayanidis, F., Andrews, S., Ward, P.B. & McConaghy, N. (1991) Effects of inter-item lag on word repetition: An event related potential study. *Psychophysiology*, 28, 307-318.

Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research,* 13, 27-39.

Logan, G.D. (1988). Towards an instance theory of automatization. *Psychological Review,* 95, 492-527.

MacLeod, C. M., & Masson, M. E. J. (2000). Repetition priming in speeded word reading: Contributions of perceptual and conceptual processing episodes. *Journal of Memory and Language, 42,* 208-228.

Masson, M. E. J. & Hicks, C, L. (1999). The influence of selection for response on repetition priming of word identification. *Canadian Journal of Experimental Psychology*, 53, 381-393.

Masson, M. E. J. & Isaak, M. I. (1999). Masked priming of words and nonwords in a naming task: Further evidence for a nonlexical basis for priming. *Memory & Cognition*, *27*, 399-412.

Masson, M.E.J. & MacLeod, C.M. (1992). Re-enacting the route to interpretation: Enhanced perceptual identification without prior perception. *Journal of Experimental Psychology: General*, *121*, 145-176.

Masson, M.E.J., & MacLeod, C.M. (2002). Covert operations: Orthographic recoding as a basis for repetition priming in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 858-871.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The Visual Word Form Area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, 7(7), 293-299.

McClelland, J. L. and Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of Basic Findings. *Psychological Review, 88,* 375-407.

McKone, E. & Trynes, K. (1999). Acquisition of novel traces in short-term implicit memory: priming for nonwords and new associations. *Memory & Cogntion,* 27, 619-632.

McKone, E. (1995). Short term implicit memory for words and nonwords. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 1108-1126.

McKone, E. (1998). The decay of short-term implicit memory: Unpacking lag. *Memory and Cognition, 26*, 1173-1186.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165-178.

Orfanidou, E., Marslen-Wilson, W. D. & Davis, M. H. (2006). Neural response suppression predicts repetition priming of spoken words and pseudowords. *Journal of Cognitive Neuroscience,* 18, 1237–1252.

Perea, M., Rosa, E. & Gómez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception & Psychophysics*, *67*, 301-314.

Pinheiro, J.C. & Bates, D.M. (2000). *Mixed-effects models in S and S-Plus.* Springer, New York.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, *12*, 767-808.

Quené, H. & van den Bergh, H. (2004). On Multi-Level Modeling of data from repeated measures designs: A tutorial, *Speech Communication*, 43 (1-2), 103-121.

Quené, H. & van den Bergh, H. (in press). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*.

R Development Core Team (2007). *R: A language and environment for statistical computing.* (version 2.4.1). [Software]. Austria: R Foundation for Statistical Computing.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with the language-as-fixed-effect fallacy: Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*, 416 – 426.

Raaijmakers, J.G.W. (2003). A further look at the "the language-as-fixed-effect fallacy". *Canadian Journal of Experimental Psychology*, *57*, 141-151.

Ratcliff, R. & McKoon, G. (1997). A counter model for implicit priming in perceptual word identification. *Psychological Review*, *104*, 319-343.

Ratcliff, R., Hockley, W. & McKoon, G. (1985). Components of activation: Repetition and priming effects in lexical decision and recognition. *Journal of Experimental* Psychology: *General, 114,* 435-450.

Raymond, J. E., Shapiro, K. L. & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 849-860

Reber, P. J., Gitelman, D. R., Parrish, T. B., & Mesulam, M-M. (2005). Priming and the acquisition of visual expertise: Changes in neural activity beyond the second presentation. *Cerebral Cortex, 15,* 787 – 795.

Reder, Angstadt Reder, L. M., Angstadt, P., Cary, M., Erickson, M. A., & Ayers, M. A. (2002a). A reexamination of stimulusfrequency effects in recognition: Two mirrors for lowand high-frequency pseudowords. Journal of Experimental Pyschology: Learning, Memory, and Cognition, 28, 138–152.

Rouder, J.N. & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin and Review*, 12, 573-604.

Rueckl, J.G. (1990) Similarity effects in word and pseudoword repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 16, 374-391.

Rumelhart, D. E., and McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The context enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60-94.

Scarborough, D.L., Cortese, C. & Scarborough, H. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 1-17.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 501-518.

Schooler, L. J., Shiffrin, R. M., Raaijmakers, J. G. W. (2001). A Bayesian model for implicit effects in perceptual identification. *Psychological Review*, *108*, 257-272.

Schütz, K., Schendzielarz, I., Zwitserlood, P. & Vorberg, D. (2007). Nice wor_ if you can get the wor_: Subliminal semantic and form priming in fragment completion. *Consciousness and Cognition*, 16, 520-532.

Shapiro, K. L.; Driver, J.; Ward, R. & Sorensen, R. E. (1997). Priming from the attentional blink: a failure to extract visual tokens but not visual types. *Psychological Science, 28, 2,* 95-100.

Shiffrin, R. M., & Steyvers, M.. A model for recognition memory: REM: Retrieving effectively from memory. *Psychonomic Bulletin and Review, 4(2)*, 145-166

Stark, C.E.L., & McClelland, J.L. (2000). Repetition priming of words, pseudowords, and nonwords. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26,* 945-972.

Swingley, D., & Aslin, R.N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*, 99-132.

Tenpenny, P.L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, 2, 339-363.

Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8*,* 336-342.

Van Turennout, M., Ellmore, T. & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience*, *3*, 1329-1334.

Venables, W. N. & Ripley, B. D. (2004). *Modern applied statistics with S.* Springer, New York.

Versace, R. & Nevers, B. (2003). Word frequency effect on repetition priming as a function of prime duration and delay between the prime and the target. *British Journal of Psychology,* 94, 389-408.

Versace, R. (1998). Frequency and prime duration effects on repetition priming and orthographic priming with words and pseudowords. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 17*, 535-556.

Visser, T. A. W., Merikle, P. M. & Di Lollo, V. (2005). Priming in the attentional blink: Perception without awareness? *Visual Cognition*, 12, 1362-1372.

Vriezen, E.R., Moscovitch, M. & Bellos, S.A. (1995). Priming effects in semantic categorization tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 933-946.

Vuilleumier, P., Schwartz, S., Duhoux, S., Dolan, R. J., & Driver, J. (2005). Selective attention modulates neural substrates of repetition priming and ''implicit'' visual memory: Suppressions and enhancements revealed by fMRI. *Journal of Cognitive Neuroscience*, 17(8), 1245–1260.

Wagenmakers, E.-J., Steyvers, M., Raaijmakers, J. G. W., Shiffrin, R. M., Van Rijn, H. & Zeelenberg, R. (2004). A model for evidence accumulation in the lexical decision task. *Cognitive Psychology*, 48, 332–367.

Wagenmakers, E.-J., Zeelenberg, R., Steyvers, M., Shiffrin, R. M., & Raaijmakers, J. G. W. (2004). Nonword repetition in lexical decision: Support for two opposing processes. *The Quarterly Journal of Experimental Psychology, 57A*, 1191–1210.

Max Planck Institute for Psycholinguistics (2001). *WebCelex.* [Internet Resource: http://celex.mpi.nl/].

Wig, G. S., Grafton, S. T., Demos, K. E. & Kelly, W. M. (2005). Reductions in neural activity underlie behavioral components of repetition priming. *Nature Neuroscience*, 8, 1228-1233.

Zeelenberg, R., Wagenmakers, E.-J. & Shiffrin, R. M. (2004). Nonword repetition priming in lexical decision reverses as a function of study task and speed stress. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 270–277.